

Appendix I – Performance Measures (Detail)

Each table includes a description of a performance measure and associated data provided by the agencies in charge of the measure. The Scope statement gives an overview of the data collection strategy for the underlying data behind the performance measure. The Source statement identifies the databases used for the measure and their proprietary agencies. The Limitations statement describes some of the shortcomings of the data in quantifying the particular performance characteristics of interest. The Statistical Issues statement are comments provided by the Bureau of Transportation Statistics (rather than the agencies in charge of the measure) that discuss variability of the measure and other points. The Verification and Validation statement indicates steps taken by the proprietary agencies to address data quality issues.

The Bureau of Transportation Statistics (BTS) feels strongly that full compliance with the Government Performance Results Act requires impartial reporting of the statistical uncertainty associated with numerical performance measures. Sometimes this assessment is straightforward, in which case we report the appropriate confidence intervals or standard errors. But sometimes assessment is not straightforward, since many sources of variation (e.g., nonsampling errors, climate effects, new technology) are difficult to quantify. Nonetheless, a combination of past data and expert judgment can enable uncertainty statements that are order-of-magnitude correct for even the most difficult problems.

The standard error of a performance measure indicates the likely size of the chance variation in the reported number. It incorporates both the effects of measurement error, survey error, and so forth, as well as the variation that occurs naturally from year to year (i.e., even if there were no change in laws, road conditions, or human behavior, there would still be chance variation in the annual count of driving fatalities---DOT success in meeting GPRA goals must be viewed in the context of this background noise). For most estimates, a 95% confidence interval on the true value of the measure includes all values within two standard errors of the reported value.

In many of the following Statistical Issues statements, BTS refers to regression standard error. This is a modification of the standard error to take account of linear trends in the recent past. Such adjustment is generally needed to incorporate consistent trends due to cumulative effects of education programs, changing demographics, the gradual adoption of new technologies, and so forth. As an example in driving safety, the combined effects of the continued educational campaign on seatbelts, the reduced percentage of drivers in the 18-22 year age range, and the broadening use of airbags are all captured as approximate linear effects. The regression standard error may be used in the same way as a regular standard error to set confidence intervals or describe uncertainty.

For further information about the source and accuracy (S&A) of these data, please refer to the BTS S&A compendium available at www.bts.gov/programs/statpol/SAcompendium.html.

Details on DOT Measures of Safety

Highway fatality and injured person rates

Page 11

Measure:	Fatalities per 100 million vehicle-miles-traveled (VMT)
Scope:	Number of fatalities come from Fatality Analysis Reporting System (FARS) data, a census of fatal traffic crashes within the 50 states, D.C. and Puerto Rico. To be included in FARS, a crash must result in the death of an occupant of a vehicle or a non-motorist within 30 days of the crash. The FARS database is a 100% count of fatal crashes collected from police crash reports, and other state data. FARS data cover all roadways open to the public, using the National Highways System classification of roads. Pedestrian and bicycle fatalities that occur on public highways but do not involve a motor vehicle are not recorded in FARS; however, this is a small number of fatalities.

	<p>Vehicle Miles of Travel (VMT) data are calculated using FHWA's Traffic Volume Trends (TVT), a monthly report based on hourly traffic count data in the Highway Performance Monitoring System (HPMS). These data, collected at approximately 4,000 continuous traffic counting locations nationwide, are used to determine the percentage change in traffic for the current month to the same month the previous year. This percentage change is applied to the travel for the same month the previous year to obtain an estimate of nationwide travel for the current month. The data are recorded as monthly totals and cumulative yearly totals.</p>
Source:	<p>NHTSA's Fatality Analysis Reporting System (FARS) for fatality data. VMT data derived from FHWA's Traffic Volume Trends (TVT) monthly report. Information is transmitted to NHTSA and entered into the system after undergoing data review by NHTSA.</p>
Limitations:	<p>VMT data are subject to sampling errors, whose magnitude depends on the how well the locations of the continuous counting locations represent nationwide traffic rates. Also, HPMS is subject to estimating differences in the states, even though FHWA works to minimize such differences and differing projections on growth, population, and economic conditions which impact driving behavior.</p>
Statistical Issues:	<p>The primary source of uncertainty in estimating fatality rates is the denominator. Although the numerator estimate of total fatalities is relatively accurate, the estimate of total vehicle miles in the denominator is far more variable. Additionally, using data from 1993-1998, the annual variation in the fatality rate attributable to random chance has a regression standard error of 0.032.</p>
Verification & Validation:	<p>Data are reviewed and analyzed by NHTSA's National Center for Statistics and Analysis. Quality control procedures are built into annual data collection at 6 and 9 months, and at year's end. A study was completed in 1993, looking at samples of FARS cases in 1989-90 to assess the accuracy of data being reported. VMT data is reviewed by FHWA for consistency and reasonableness.</p>
Comment:	<p>The data programs have been around for many years and are generally accepted for describing safety on the Nation's highways. Adjusting raw highway fatalities and injuries by VMT provides a means of portraying the changes in highway fatalities on a constant exposure basis – to facilitate year-to-year comparisons.</p>

Highway fatality and injured person rates

Page 11

Measure:	Injured persons per 100 million vehicle-miles-traveled (VMT)
Scope:	<p>Injured persons data are derived from General Estimates System (GES), a nationally representative probability sample that makes national estimates of total nonfatal injury crashes, injured persons, and property-damage-only crashes. GES data cover all roadways open to the public, using the National Highways System classification of roads.</p> <p>Vehicle Miles of Travel (VMT) data are calculated using FHWA's Traffic Volume Trends (TVT), a monthly report based on hourly traffic count data in the Highway Performance Monitoring System (HPMS). These data, collected at approximately 4,000 continuous traffic counting locations nationwide, are used to determine the percentage change in traffic for the current month to the same month the previous year. This percentage change is applied to the travel for the same month the previous year to obtain an estimate of nationwide travel for the current month. The data are collected as monthly totals and cumulative yearly totals.</p>
Source:	<p>NHTSA's General Estimates System (GES) provides injury data. VMT data are derived from FHWA's Traffic Volume Trends (TVT) monthly report.</p>
Limitations:	<p>GES data are obtained from a nationally representative sample of 60 sites. The results provide only national data, not state by state data. VMT data are subject to sampling errors, whose magnitude depends on the how well the locations of the continuous counting locations represent nationwide traffic rates. Also, HPMS is subject to estimating differences in the states, even though FHWA works to minimize such differences and differing projections on growth, population, and economic conditions which impact driving behavior.</p>
Statistical Issues:	<p>The estimate of the injury rate includes three main sources of uncertainty. The numerator count of injuries has a standard error of 7.9% (cf. Appendix C of <i>Traffic Safety Facts</i>). The denominator estimate of VMT contains complex sampling and nonsampling error. Additionally, chance variation within a year has, based on data from 1993-1998, a regression standard error of 6.32.</p>

Verification & Validation:	Data are reviewed and analyzed by NHTSA's National Center for Statistics and Analysis. Quality control procedures are built into annual data collection at 6 and 9 months, and at year's end. VMT data are reviewed by FHWA for consistency and reasonableness.
Comment:	The data programs have been around for many years and are generally accepted for describing safety on the Nation's highways. GES records injury severity in four classes: incapacitating injury, evident injury but not incapacitating, possible but not visible injury, and injury of unknown severity. Adjusting raw highway fatalities and injuries by VMT provides a means of portraying the changes in highway fatalities on a constant exposure basis – to facilitate year-to-year comparisons.

Alcohol related highway fatalities

Page 13

Measure:	Percentage of highway fatalities that are alcohol related
Scope:	The Fatality Analysis Reporting System (FARS) database is a census of fatal crashes within the 50 states, D.C. and Puerto Rico. FARS data cover all roadways open to the public, using the National Highways System classification of roads. To be included in FARS, a crash must result in the death of a vehicle occupant \or a non-motorist within 30 days of the crash. A fatal crash is alcohol-related if either a driver or a non-motorist (such as a pedestrian) had a measurable or estimated blood alcohol concentration (BAC) of 0.01 grams per deciliter or above.
Source:	NHTSA's Fatality Analysis Reporting System (FARS).
Limitations:	Blood Alcohol Concentration (BAC) test results are not available for all drivers and non-occupants involved in fatal crashes. Missing data can result for a number of reasons; the most frequent of which is that persons are not always tested for alcohol. To address the missing data issue, NHTSA has developed and employed a statistical model to estimate the likelihood that a fatal crash-involved a driver who was sober (BAC of zero), had some alcohol (BAC of 0.01-0.09), or was intoxicated (BAC of 0.10 or greater) at the time of the crash. The statistical model (applied since 1982) is based on important characteristics of the crash including crash factors, vehicle factors, and person factors. While this measure does not link alcohol with fault in fatal crashes, the more comprehensive scope of the measure compensates for a possible undercount of the size of the alcohol impaired driving problem.
Statistical Issues:	The primary sources of uncertainty in this performance measure arise from information gaps in the number of intoxicated non-motorists, and from using the statistical model to estimate the number of intoxicated drivers. Chance variation within the year, based on 1993-1998 data, has a regression standard error of 0.86.
Verification & Validation:	Data are reviewed and analyzed by NHTSA's National Center for Statistics and Analysis. Quality control procedures are built into annual data collection at 6 and 9 months, and at year's end. In 1988, an independent panel of academics reviewed and commented on the statistical methods used in measuring alcohol-related highway fatalities. This report supported the approach currently in use.
Comment:	The data program has been used for many years and is generally accepted for describing safety on the Nation's highways. In 2000, this performance measure was revised to reflect the percentage of highway fatalities that are alcohol related. NHTSA believes that percentage targets are better annual measures because they factor in the overall traffic fatality number and can be predicted with greater precision than total numbers of alcohol-related fatalities.

Seat belt use**Page 15**

Measure:	Percent of front seat occupants using seat belts.
Scope:	Based on the National Occupant Protection Use Survey (NOPUS). NOPUS is a National multi-stage, probability sample. In the first stage, counties were grouped by region (Northeast, Midwest, South, West), level of urbanization (metropolitan or not), and level of belt use (high, medium, or low). Fifty counties or groups of counties were selected based on the vehicle miles of travel in those locations. In the next stage, a random sample of 4,000 roadway segments for all types of roads was selected. In even numbered years observation of shoulder belt use is conducted at all 4,000 of these roadway segments. In the 1999 survey, a mini-NOPUS was conducted, consisting of 2,000 of the 4,000 segments chosen for observation. Observation includes driver and right front seat passenger in passenger cars, pickup trucks, vans, and sport utility vehicles.
Source:	Actual Observation of shoulder belt use in moving vehicles by teams of observers. Belt use was observed only in daylight hours.
Limitations:	Data are based on a random sample of sites and, therefore, are subject to sampling error. Sampling errors for the 1999 survey were estimated to be approximately 1.3 percentage points.
Statistical Issues:	The primary source of uncertainty in the ultimate calculation of this measure derives from disparities among the surveys and the use of non-probability samples for some of them. The GPRA report is based entirely upon the NOPUS probability sample (see the Verification and Validation section), but the survey bias and reweighting are complex. The regression standard error in the reported percentage, based on 1993-1998 data, is approximately 0.37.
Verification & Validation:	Data collection is managed by a survey research contractor who has responsibility to hire and train the data collectors/observers. NHTSA reviews and approves all the training materials. Data collectors/observers must pass a 2-day training course in obtaining the data before data collection begins. The data contractor also conducts on-scene "surprise" quality control visits to ensure that observations are made correctly and data are coded properly. Numerous edits are also employed in the data processing. NHTSA reviews the data provided by the contractor for consistency.
Comment:	None.

Large truck-related fatalities and injured persons**Page 17**

Measure:	Number of fatalities involving large trucks.
Scope:	The measure includes all fatalities (occupants of passenger cars, motorcycles, large trucks, or pedestrians, etc.) associated with crashes involving trucks with a gross vehicle weight rating of 10,000 pounds or more. Number of fatalities come from Fatality Analysis Reporting System (FARS) data, a census of fatal traffic crashes within the 50 states, D.C. and Puerto Rico.
Source:	NHTSA's Fatality Analysis Reporting System (FARS) provides fatality data.
Limitations:	FARS data elements are modified from year to year to respond to emphasis areas, vehicle fleet changes, and other needs for improvement.
Statistical Issues:	The fatality counts in FARS are generally accurate---the major sources of error are underreporting by some precincts and inconsistent use of the definition of a truck. Based on 1993-1998 data, the chance variation in a given year has a regression standard error of 130.9.
Verification & Validation:	Fatality data are reviewed and analyzed by NHTSA's National Center for Statistics and Analysis. Quality control procedures are built into data collection and data processing. A study was completed in 1993, looking at samples of FARS cases in 1989-90 to assess the accuracy of data being reported.
Comment:	The data have been around for many years and are generally accepted as a good source for describing fatal crashes on the Nation's highways.

Large truck-related fatalities and injured persons

Page 17

Measure:	Number of injured persons involving large trucks.
Scope:	The measure includes all injured persons (occupants of passenger cars, motorcycles, large trucks, or pedestrians, etc.) associated with crashes involving trucks with a gross vehicle weight rating of 10,000 pounds or more. Injury data are derived from General Estimates System (GES).
Source:	NHTSA's General Estimates System (GES) provides injury data.
Limitations:	GES data are obtained from a nationally representative sample of 60 sites. The results provide only national data, not state-by-state data.
Statistical Issues:	The GES data have a standard error of 7.9% for injuries from truck and automobile crashes (cf. Appendix C of <i>Traffic Accident Reports</i>), and are less accurate than the corresponding fatality counts. Based on 1993-1998 data, the variation due to random chance in these numbers, which includes sampling variability, has a regression standard error of approximately 6,940.
Verification & Validation:	Injury data are reviewed and analyzed by NHTSA's National Center for Statistics and Analysis. Quality control procedures are built into data collection and data processing.
Comment:	The data program has been around for many years and is generally accepted for describing safety on the Nation's highways. GES records injury severity in four classes: incapacitating injury, evident injury but not incapacitating, possible but not visible injury, and injury of unknown severity.

Air carrier fatal accident rate

Page 19

Measure:	Fatal aviation accidents (commercial air carriers) per 100,000 flight hours.
Scope:	This measure includes both scheduled and nonscheduled flights of large U.S. air carriers (FAR Part 121) and scheduled flights of commuter airlines (FAR Part 135). It excludes on-demand (i.e., air taxi) service and general aviation.
Source:	Part 121 and Part 135 flight hour data is submitted to BTS under FAR Parts 241 and 298, respectively. NTSB provides accident data.
Limitations:	The fatal accident rate in these categories is small and could significantly fluctuate from year to year due to the occurrence or non-occurrence of a single accident.
Statistical Issues:	There is no significant error in either the accident counts (numerator) or the flight hours data (denominator) for this rate. However, random variation in air crashes results in a significant variation in the accident rate from year to year. The regression standard error in this variation for 1993 through 1998 is .017. Therefore, a 90 percent confidence interval for the 1999 accident rate of .04 is [.012, .068].
Verification & Validation:	The FAA does comparison checking of the flight hours reported to BTS with hours reported on the Air Carrier Utilization Reports. BTS uses clerical and computer edit and range checks to verify that data are correct. NTSB and FAA's Office of Accident Investigation meet regularly to validate the accident count.
Comment:	This goal assumes a 15 % reduction in fatal accidents in the following five areas covered by <i>Safer Skies – A Focused Agenda</i> : controlled flight into terrain, loss of control, uncontained engine failure, approach and landing, and weather. The sixth area in <i>Safer Skies</i> , runway incursions, is the subject of a separate performance goal. These causal factors accounted for 14 of the 18 total fatal accidents in the baseline years 1994 through 1996. The net reduction – about 12% – reflects a 15% reduction in areas that cover about 78% of the accidents.

General aviation fatal accidents**Page 21**

Measure:	Number of general aviation fatal accidents.
Scope:	The measure includes on-demand (non-scheduled FAR Part 135) and general aviation fatal accidents. General aviation comprises a diverse range of aviation activities. The range of general aviation aircraft include single-seat homebuilt aircraft, helicopters, balloons, single and multiple engine land and seaplanes including highly sophisticated extended range turbojets.
Source:	National Transportation Safety Board
Limitations:	The use of the most recent three-year period for the baseline represents one of the safest periods in general aviation history in terms of a decline in fatal accidents.
Statistical Issues:	Like the air carrier data described above, there is no significant error in the accident counts. Random variation in air crashes results in a significant variation in the number of fatal accidents over time. The regression standard error in this variation for 1993 through 1998 is 18.7. Therefore, a 90 percent confidence interval for the 1999 accident number of 354 is [323.2, 384.8].
Verification & Validation:	NTSB and FAA's Office of Accident Investigation meet regularly to validate the information on the number of accidents.
Comment:	It would be preferable to use fatal accident rates rather than fatal accidents as the performance measure. However, general aviation flight hours are based on an annual survey conducted by the FAA. Response to the survey is voluntary. The accuracy of the flight hours collected is suspect and there is no readily available way to verify or validate the data. For this reason, the General Aviation community is unwilling to use a rate measure until the validity and reliability of the survey data can be assured.

Runway incursions**Page 23**

Measure:	Number of runway incursions.
Scope:	Incursions occur as a result of operational errors, operational deviations, surface pilot deviations, or vehicle/pedestrian deviations. Incursions are reported and tracked at airports that have an operational air traffic control tower.
Source:	Air traffic controllers and pilots are the primary source of runway incursion reports. The data is recorded in the FAA National Incident Monitoring System (NAIMS).
Limitations:	Preliminary incident reports are evaluated when received. Evaluation can take up to 90 days.
Statistical Issues:	Sources of nonsampling error, such as human error in completing reports have not been quantified to date.
Verification & Validation:	Surface incidents are reported in the Administrator's Daily Bulletin at the beginning of each weekday. Surface incidents are evaluated to determine if they should be classified as incursions. Incidents are evaluated against the official runway incursion definition. The Air Traffic Runway Safety Program Manager, ATP-20, makes the final decision regarding runway incursions.

Operational Errors (Air Traffic)**Page 26**

Measure:	1. Operational errors per 100,000 facility activities. 2. Deviations per 100,000 facility activities.
Scope:	An error occurs when separation between aircraft is less than the separation determined necessary for the specific phase of flight. An operational deviation occurs when an aircraft enters airspace without prior coordination. "Activities" are total facility activities, as defined in <i>Aviation System Indicators 1997 Annual Report</i> . Total facility activities are the sum of en route and terminal facility activities.
Source:	FAA air traffic facilities have a software program called Operational Error Detection Patch (OEDP) that detects possible operational errors and sends alert messages to supervisory personnel. Facility management review OEDP alerts and data provided from the National Track Analysis Program (NTAP) to determine if an operational error has occurred. Controllers are required to report both operational errors and operational deviations. The information is summarized in the FAA Air Traffic Operational Error and Deviation Database.
Limitations:	<p>There is a few months' lag in reporting data because of the need to investigate significant incidents. The severity of errors is not measured. Minor errors such as a 4 and 1/2 mile rather than a 5-mile separation are counted in the same way as more serious errors. Data are available for 1994 and following years.</p> <p>The Inspector General (IG) is currently conducting an audit of reporting on operational errors. The IG believes that there is a potential for underreporting of operational errors as some errors are self-reported. The FAA disagrees with this assessment because there are significant penalties for not reporting an operational error.</p>
Statistical Issues:	Like the runway incursion data described above, there are no significant sources of systematic error in the operational errors and deviations data that have been quantified. Again, random variation in operational errors and deviations results in a significant variation in the measured rates over time.
Verification & Validation:	FAA performs system checks and counts daily against reported data to ensure the accuracy of information reported.
Comment:	In August 1998, the FAA discovered and corrected a misunderstanding of the procedures used in interpreting separation reported by the National Track Analysis Program and the data provided by the Operational Error Detection Patch. The corrected application of these procedures while not affecting safety, has resulted in overall increase in the number of errors report between 4.6 and 4.9 miles separation (Standard separation in these cases is 5 miles).

Recreational boating fatalities**Page 28**

Measure:	Number of recreational boating fatalities
Scope:	Measure includes fatalities occurring aboard vessels that are being operated for recreational purposes. Surfboards, iceboats, and vessels engaged in sanctioned racing events are not considered recreational vessels. Fatalities are included if caused by a fire, explosion, sinking or other occurrence involving a recreational vessel, and the vessel or associated equipment caused or contributed to the fatality. Fatalities are not included if they occurred aboard a recreational vessel, but were caused by self-inflicted wounds or natural causes. Fatalities are also excluded if they occurred while the victim was engaged in other activity such as swimming or diving, where the vessel was used as a platform only and was not a contributing factor to the fatality. Beginning this year, the measure for Recreational Boating will be revised to account for an estimated 6% underreporting of recreational boating fatalities.

Source:	Coast Guard Boating Accident Report Database (BARD). Data is entered into BARD by state administrators who collect data from boat owners and operators through formal Boating Accident Reports, as instructed in 33 CFR 173c.
Limitations:	Fatality data is derived from reports submitted by the public along with accompanying state investigation reports. There is consensus among the Coast Guard, the states, safety professionals, and other researchers that most fatalities that occur on inland and most coastal waters are reported. To better quantify the extent of possible under-reporting the Coast Guard initiated and funded an analysis of BARD data conducted by the Boat Owners Association of the United States (BOAT/U.S.) Foundation for Boating Safety. The study found some fatalities involving recreational boating in the Coast Guard's Search and Rescue Management Information System (SARMIS) that were not in BARD. However, although the study reported a 9% discrepancy, further analysis revealed that some of these findings would not be reportable as recreational boating fatalities. More analysis is underway to confirm the true size of this discrepancy. There is also consensus that under-reporting exists for fatalities occurring offshore, and aboard U.S. recreational boats operating overseas. Also, although there are guidelines as to what constitutes a recreational boating fatality, there is still an element of interpretation at the state level in reporting fatalities. It is probable that the states do not always interpret the guidelines in the same manner. Overall, the best estimate indicates that total fatalities are currently under-reported by at least 6%.
Statistical Issues:	The discrepancy between BARD and the Search & Rescue Management Information System (SARMIS) amounts to 4% of the total reports for those states covered by SARMIS. The numbers given in this report have been adjusted to correct the deficiency. Also, note that the boating fatality counts are strongly influenced by weather; annual chance variation should be large.
Verification & Validation:	Fatality data in BARD is verified and validated by state boating administrators and Coast Guard program managers. At the end of the calendar year, the Coast Guard compiles state fatality data and sends a report to each state for confirmation. State and Coast Guard officials both review the statistics, including sampling of cases to ensure guidelines for classifying fatalities were followed. Any discrepancy is reconciled jointly by the state and Coast Guard program manager.
Comment:	<p>Data are not normalized for increases or decreases in the number or usage of boats, which tend to limit data use in making comparisons over time. The number and usage of recreational boats has increased over the past 2 decades, while the raw number of fatalities has generally decreased.</p> <p>The BOAT/US review of BARD data for 1993 through 1997 identified underreporting in BARD of 8% in 1993 and 1994, 12% in 1995, 13% in 1996 and 8% in 1997. The Coast Guard reviewed BOAT/US's findings for 1995, 1996, and 1997. Each record for these years was checked and fatalities that were incorrectly labeled as recreational boating fatalities by BOAT/US were removed from the count. Based on this revised count of recreational boating fatalities with mislabeled fatalities removed, the Coast Guard estimates that 7%, 8% and 4% of all recreational boating fatalities were not captured in its Boating Accident Report Database (BARD) in 1995, 1996 and 1997 respectively for purposes of this report. The median of these numbers – 6% - has been used to adjust recreational boating safety data for 1993, 1994, 1998 and 1999, and to reset the goals for 1999 through 2001. The original goal of 720 has been increased by 6% to 763.</p> <p>The Coast Guard is in the process of commissioning a comprehensive National Boating Survey to obtain valid and reliable information on boating practices, safety, and exposure. This information will enable safety officials to assess boating risk, implement appropriate safety intervention strategies, and measure the effectiveness of program activities in reducing the risk and negative outcomes associated with the use of recreational boats. Data from this study will be used to further address underreporting issues and estimate reporting discrepancies in BARD but will not be available for over two years.</p>

Maritime search and rescue - (FY 1999 – FY 2000)

Page 30

Measure:	1. Percent of mariners reported in imminent danger who are rescued. 2. Percent of property reported in imminent danger saved.
Scope:	<p>For mariners, the measure includes people in water; on shore; aboard a vessel, offshore structure, pier, or vehicle that is reported to the Coast Guard as being in distress or in urgent need of assistance. For property, the measure indicates value of property that is reported to the Coast Guard as being in imminent danger of being lost. The Coast Guard makes a final determination on scene whether there is imminent danger, based on criteria that take into account the nature of distress, condition of the vessel, the people onboard, and the environmental conditions. Criteria for this decision are discussed in search and rescue doctrine publications. Value of property is determined through means including sales, financial and insurance records, comparisons to similar property, and owner/operator input.</p>
Source:	<p>Coast Guard Search and Rescue Management Information System (SARMIS). Data is collected from Coast Guard field units that conduct search and rescue responses.</p>
Limitations:	<p>It is probable that some number of imminent danger cases, and the associated lives and property, are not reported in SARMIS. This includes situations where no distress call was received by the Coast Guard and the persons in distress were rescued by private citizens or local government personnel, or where the persons in distress perished without trace. The extent of this under-reporting is not known. There is also some judgment involved in assessing whether mariners or property are in imminent danger. Judgment is also involved in assessing property value where assessed values are not available. However, there is likely to be consistency in these assessments across years. 1994 data is skewed upward by a large surge of migrants interdicted at sea, most of whom were counted as "rescued," thus increasing the percentage of lives and property reported as saved. Reporting no longer includes migrants interdicted; they are counted directly as migrants interdicted under law enforcement activity. Errors may be introduced in SARMIS through data entry, but are likely rare for lives/property saved data elements.</p>
Statistical Issues:	<p>The primary source of uncertainty consists of non-sampling errors. The SARMIS system runs the risk of missing cases or duplicating cases, since case numbers are not machine-generated. SARMIS II (which will be in use beginning with FY 2000) will assign machine generated case numbers. SARMIS II will also reduce the error due to miscoding. The regression standard error for year-to-year chance variation in the measures are 1.53 for mariners and 10.3 for property, so that a 90% confidence interval for the 1999 values of the measures, 95% and 80% respectively, are [92.5, 97.5] and [63.0, 87.0] respectively.</p>
Verification & Validation:	<p>SARMIS data entry system uses structured entry values, check boxes, and pull down selection lists to limit entry errors. The use of plain language descriptions eliminates a majority of erroneous data code selection. Additional system business rules also eliminate the selection of data not appropriate with other entered data. The SAR Mission Coordinator (SMC) is responsible for accurate entry of particular case data by all units involved in the case. CG Program Managers annually validate the data in SARMIS. Entries are reviewed at Coast Guard District offices as first step in validation – errors and inconsistencies are identified and corrected. Finally Coast Guard headquarters program managers review compiled data annually to assess consistency with historic variance and trends. This review includes: curvilinear regression analysis to compare current data to historic data, and a program review analysis to identify and resolve aberrations.</p>
Comment:	<p>After FY00, the measure will be changed from all mariners reported in distress to all mariners in distress.</p>

Maritime search and rescue - (FY 2001)**Page 30**

Measure:	Percent of all mariners in imminent danger who are rescued.
Scope:	Includes people in water; on shore; aboard a vessel, offshore structure, pier, or vehicle that is in distress or in urgent need of assistance. The Coast Guard makes a final determination on scene whether there is imminent danger, based on criteria that include the nature of distress, the condition of the vessel, the people onboard, and the environmental conditions. Criteria for this decision are discussed in search and rescue doctrine publications.
Source:	CG Search and Rescue Management Information System (SARMIS). Data is collected from Coast Guard field units that conduct search and rescue responses.
Limitations:	It is probable that some number of imminent danger cases, and the associated lives and property, are not reported in SARMIS. This includes situations where no distress call was received by the Coast Guard and the persons in distress were rescued by private citizens or local government personnel, or where the persons in distress perished without trace. The extent of this under-reporting is not known. There is some judgment involved in assessing whether mariners are in danger. However, there is likely to be consistency in these assessments across years. 1994 data is skewed upward by a large surge of migrants interdicted at sea, most of whom were counted as "rescued," thus increasing the percentage of lives reported as saved. Reporting no longer includes migrants interdicted; they are counted directly as migrants interdicted under law enforcement activity. Errors may be introduced in SARMIS through data entry, but are likely rare for lives/property saved data elements.
Statistical Issues:	The primary source of uncertainty consists of non-sampling errors. The SARMIS system runs the risk of missing cases or duplicating cases, since case numbers are not machine-generated. SARMIS II (which will be in use beginning with FY 2000) will assign machine generated case numbers. SARMIS II will also reduce the error due to miscoding.
Verification & Validation:	SARMIS data entry system uses structured entry values, check boxes, and pull down selection lists to limit entry errors. The use of plain language descriptions eliminates a majority of erroneous data code selection. Additional system business rules also eliminate the selection of data not appropriate with other entered data. The SAR Mission Coordinator (SMC) is responsible for accurate entry of particular case data by all units involved in the case. CG Program Managers annually validate the data in SARMIS. Entries are reviewed at Coast Guard District offices as first step in validation – errors and inconsistencies are identified and corrected. Finally Coast Guard headquarters program managers review compiled data annually to assess consistency with historic variance and trends. This review includes: curvilinear regression analysis to compare current data to historic data, and a program review analysis to identify and resolve aberrations.
Comment:	Beginning in FY01, this measure will cover all mariners in distress. The previous measure covered only mariners reported in distress that were rescued. The significance of the 88% result for FY99 is uncertain at this point; FY95-98 data show a flat trend at 84%. It is not known if the FY99 result was produced by anomalous factors, or if it is the product of program strategies and a changing external environment. Therefore the goal target remains at 85% until more analysis is completed.

Passenger Vessel Safety**Page 32**

Measure:	Number of high-risk passenger vessel casualties per 1000 vessels.
Scope:	Included all commercial passenger vessels carrying more than 6 passengers for hire operating in U.S. navigable waters are included, as well as U.S. passenger vessels operating in any waters, that are investigated by Coast Guard personnel. High-risk passenger vessel casualties include fire, explosion, capsizing, flooding, allision, collision, capsizing, grounding, or sinking.

Source:	Casualties (numerator) are obtained from the Coast Guard Marine Safety Information System (MSIS). Passenger vessel casualties are reported to the Coast Guard by vessel operators, owners, insurance companies, legal representatives, and other mariners. The total number of passenger vessels (denominator) is a Coast Guard count of all inspected passenger vessels embarking passengers in U.S. ports.
Limitations:	1999 data are preliminary estimates base on approximately 7 months of data. Data was extrapolated for a full year based on historical data patterns. It is probable that some high-risk casualties are not reported. This number is unknown. During the last 8 years, the Coast Guard has initiated violation cases against 351 entities for failure to report marine casualties, although many of these were for minor casualties. Duplicate casualty entries are sometimes entered into MSIS, and some casualties are mistakenly omitted or coded incorrectly. Verification procedures strive to correct these errors, but it is probable that a small number are not corrected. It is also possible that errors exist in the data for number of passenger vessels, although this error, if it exists, is probably extremely low. This measure is an indirect indicator of the risk of major loss of life, and serves as a leading indicator of passenger vessel safety.
Statistical Issues:	The primary source of uncertainty in estimating the performance measure is probably the response error. While there is a strong penalty for not reporting fatalities, the Coast Guard has initiated violation cases against 351 individuals and companies during the last 8 years for failure to report marine casualty cases. The number of unreported vessel casualties is greater than or equal to the number detected for violations.
Verification & Validation:	Verification and validation occurs at several levels. Edit checks within MSIS can detect some incorrect or missing data and force review and correction before data entry is completed. Selection lists for certain data fields also reduce the opportunity for data entry error. All investigations go through one level of review at the field unit for accuracy. Investigations of serious marine casualties are also usually reviewed at district and headquarters offices. The headquarters Data Administration staff conducts periodic quality control checks to identify entry errors such as missing data or miscoding, and corrects any errors identified. Each investigation involving the loss of a vessel is reviewed before it is included in the measure. Errors identified are referred to either the Data Administration staff or the Investigations and Analysis staff for correction.
Comment:	It is important to remember that this is not a measure of deaths on passenger vessels. Rather, it measures events (passenger vessel accidents) that increase the risk of a major loss of life on passenger vessels.

Maritime Worker Fatality Rate

Page 34

Measure:	Fatalities per 100,000 workers aboard commercial vessels.
Scope:	Includes fatalities aboard commercial vessels and marine platforms within U.S. waters, or on board U.S. flag commercial vessels anywhere in the world, that are investigated by Coast Guard personnel. Does not include fatalities involving non-workers, recreational vessels, foreign vessels, homicides, suicides, or natural causes.
Source:	Fatalities (numerator) are obtained from the Coast Guard Marine Safety Information System (MSIS). Data is collected from field units that investigate fatality cases. Worker population estimates (denominator) are obtained from National Marine Fisheries Service, Bureau of Labor Statistics, and Mineral Management Service. Categories of workers include fishing vessels, towboats, freight/tankships, passenger vessels, offshore drilling units, outer continental shelf platforms, and offshore supply vessels.
Limitations:	1999 data are preliminary estimates based on approximately 7 months of data. Data was extrapolated for a full year based on historical data patterns. It is probable that some fatalities are not reported. This number is probably extremely low. Duplicate fatality entries are sometimes entered into MSIS, and some fatalities are mistakenly omitted or coded incorrectly. Verification procedures strive to correct these errors, but it is probable that a small number are not corrected. Worker population data is collected by other agencies through sampling and survey techniques that are subject to coverage and response rate errors. Collection methodology revisions also affect a year-to-year comparison of population estimates. Current year population estimates are not always available, so a 5-year weighted average is used. The Bureau of Labor Statistics indicates that substantial estimating error exists, particularly in the fishing industry population. Using worker population as a normalizing factor is a surrogate for quantifying risk. There may be conceptually better normalizing factors, but worker population appears to be the best factor currently available.
Statistical Issues:	The primary source of uncertainty in estimating the performance measure is probably the response error. While there is a strong penalty for not reporting fatalities, the Coast Guard has initiated violation cases against

Verification & Validation:	351 individuals and companies during the last 8 years for failure to report marine casualty cases. The number of unreported vessel casualties is greater than or equal to the number detected for violations. Verification and validation occurs at several levels. Edit checks within MSIS can detect some incorrect or missing data and force review and correction before data entry is completed. Selection lists for certain data fields also reduce the opportunity for data entry error. All investigations go through one level of review at the field unit for accuracy. Investigations of fatalities are also usually reviewed at district and headquarters offices. The headquarters Data Administration staff conducts periodic quality control checks to identify entry errors such as missing data or miscoding, and corrects any errors identified. Each investigation involving a fatality is reviewed before it is included in the measure. Errors identified are referred to either the Data Administration staff or the Investigations and Analysis staff for correction.
Comment:	None.

Rail accident and fatality rates

Page 35

Measure:	1. Train accidents per million train-miles. 2. Rail-related fatalities per million train-miles.
Scope:	The fatality measure includes anyone on rail property, any on-duty railroad employee, and anyone killed by a train or its contents. It does not include fatalities on trains or rail lines that do not connect to the national rail network, such as mass transit operations, certain excursion and tourist railroads, and some industrial railroads not connected to the general system. The only railroad fatalities that are not counted are suicides (as determined by a public official) and death by natural cause not associated with railroad operations.
Source:	<i>Railroad Safety Statistics – Annual Report</i> . Statistical data, tables, and charts depict the causes and nature of rail-related fatalities and accidents. Data on fatalities, accidents, and train miles are reported to FRA by railroad companies.
Limitations:	Because of the scope of the reporting criteria, some fatalities that are counted are not associated directly with operation of the trains, and some railroad fatalities are not counted. This scope is consistent with the regulatory authority of the agency, but not consistent with other modes of transportation for comparative purposes.
Statistical Issues:	The reported estimates are based upon the first 11 months of data from 1999. Totals were adjusted to an annual rate by a correction factor. Chance variation from year to year, as reflected in the regression standard error, is 0.199 for train accidents and 0.079 for rail fatalities.
Verification & Validation:	Railroads are required by law to submit monthly accident/incident reports to FRA. They are also required to update any inaccurate or incomplete information. FRA conducts routine data audits (records inspections) to verify the adequacy of railroad reporting and recordkeeping requirements.
Comment:	None.

Highway - Rail grade-crossing accidents**Page 37**

Measure:	Grade-crossing accidents divided by the product of: 1) million train miles and 2) trillion vehicle miles of travel.
Scope:	The measure includes all collisions with on track equipment and highway users at public and private grade crossings.
Source:	Collisions and train-miles are reported in FRA's <i>Railroad Safety Statistics – Annual Report</i> . Vehicle-miles-traveled (VMT) are obtained from the FHWA Office of Highway Information Management.
Limitations:	Because the denominator includes all highway vehicle-miles-traveled (VMT), and not just VMT that are exposed to grade crossings, the rate portrayed may be lower than the actual risk.
Statistical Issues:	Trains and automobiles have different exposures at rail crossings---the denominator used here attempts to combine these. The numerator is based on the first 11 months of 1999 and an adjustment was used to obtain the estimated rate for the year. The annual variation by chance from year to year as measured by the regression standard error is 0.022.
Verification & Validation:	FRA's Office of Safety has a review process to ensure that railroads and the States comply with Federal reporting requirements in the preparation of the FRA <i>Railroad Safety Statistics - Annual Report</i> .
Comment:	The measure is a ratio of total highway-rail grade crossing collisions, total-train miles, and total highway vehicle-miles-traveled. The targets reflect the trend necessary to meet FRA's out-year goal of a 1.23 rate in 2002.

Rail trespasser fatality rate**Page 39**

Measure:	Number of rail-related trespasser fatalities divided by the product of: 1) million train-miles and 2) billion U.S. population.
Scope:	Trespassers are persons who are on that part of railroad property used in railroad operations and whose presence is prohibited, forbidden or unlawful.
Source:	Trespasser fatalities are reported in FRA's <i>Railroad Safety Statistics - Annual Report</i> using data submitted by railroads to FRA's Office of Safety. The U.S. population data are from the U.S. Bureau of the Census.
Limitations:	The denominator reflects an attempt to capture changes in risk exposure associated with increasing train mileage, coupled with increasing population that may potentially trespass on railroad property. Because not all of the population is exposed to railroads, the rate portrayed may be lower than the actual risk.
Statistical Issues:	The numerator is based on the first 11 months of 1999 and an adjustment was used to obtain the estimated rate for the year. The product in the denominator may not provide an interpretable measure of exposure. The identification of suicides may not be commensurate across localities. The annual chance variation as measured by regression standard error in this estimate is 0.189.
Verification & Validation:	Railroads are required by law to submit monthly trespasser fatality reports to FRA. They are also required to update any inaccurate or incomplete information. FRA conducts routine data audits (records inspections) to verify the adequacy of railroad reporting and record keeping requirements.
Comment:	Suicides as determined by a public official are not counted.

Transit fatality and injured person rates**Page 40**

Measure:	1. Transit fatalities per 100 million passenger miles traveled. 2. Transit injured persons per 100 million passenger miles traveled.
Scope:	The data include both riders and employees. A fatality is defined as a transit-caused death from collision, personal casualty, fire, derailment, or bus going off the road. An injury is defined as any physical damage or harm to a person requiring medical treatment caused by a transit collision, personal casualty, fire, derailment, or bus going off the road.
Source:	FTA's Safety Management Information System (SAMIS), with data reported by transit operators to the National Transit Database (NTB).
Limitations:	Because of the scope of the reporting criteria, some fatalities that are counted are not associated directly with transit operation. This scope is consistent with the regulatory authority of the agency, but not consistent with other modes of transportation for comparative purposes.
Statistical Issues:	The fatality and injury counts in SAMIS are generally quite accurate---the major source of error in the measure comes from uncertainty in the passenger miles traveled. Based on 1993-1998 data, the chance variation in a given year has a regression standard error of 0.045 for the transit fatality rates and 4.617 for the transit injury rates.
Verification & Validation:	An independent auditor and the transit agency's CEO certify that data reported to the NTD are accurate. Using data from the NTD to compile the SAMIS data, the Transportation Systems Center compares current safety statistics with previous years, identifies questionable trends, and seeks explanation from operators.

Pipeline failures**Page 42**

Measure:	1. Failures of natural gas transmission pipelines. 2. Failures of hazardous liquid pipelines.
Scope:	This measure is based on reported hazardous natural gas leaks that meet federal reporting criteria as defined in 49 CFR 191.1 and 191.15 for natural gas transmission pipeline incidents.
Source:	RSPA's Natural Gas Transmission Incident Report and Hazardous Liquid Accident Report. Failure reports are filed within 30 days of the occurrence of reportable incidents. Complete calendar year data are available by March 1 of the following year. Data may change as operators file supplemental reports.
Limitations:	RSPA lacks adequate infrastructure information on pipeline operations and maintenance needed to fully characterize problems when they occur and lacks information on precursor conditions that contribute to incidents. Joint Federal, state and industry teams have been formed to devise a new course to improve information availability.
Statistical Issues:	The number of failures for both natural gas and hazardous liquid are likely to be underreported. The annual variation in the number of failures from year to year due to chance has a regression standard error of 252 for natural gas and 15.1 for liquid pipeline based on data from 1993 to 1998.
Verification & Validation:	RSPA reviews/verifies data provided for accuracy and requests supplemental reports where shortcomings are indicated.
Comment:	Hazardous liquid pipeline spills are not included as a safety goal, as they primarily impact the environment.

Pipeline failures**Page 42**

Measure:	Number of pipeline incidents caused by outside force damage.
-----------------	---

Scope:	Natural gas pipeline incidents are those that result in a fatality or injury resulting in hospital treatment or property damage equal to or greater than \$50,000. Hazardous liquid pipeline incidents are those that result in a fatality or injury resulting in medical treatment or property damage equal to or greater than \$50,000 or over 50 barrels spilled.
Source:	RSPA's natural gas distribution and transmission pipeline incident reports and RSPA's hazardous liquid accident reports.
Limitations:	RSPA lacks adequate infrastructure information on pipeline operations and maintenance needed to fully characterize problems when they occur and lacks information on precursor conditions that contribute to incidents. Joint Federal, state and industry teams are working to improve data collection and usefulness. Excavation damage is a leading cause of pipeline failures, due to failure to call for location information before digging, incorrect marking of locations, and failure to observe markings. RSPA is studying state program damage prevention initiatives to determine best practices and is working to help educate the public about the need to call before digging.
Statistical Issues:	Although the number of incidents is likely to be underreported, such recording error is probably small in comparison to the annual variation. The annual variation in the number of failures from year to year due to chance has a regression standard error of 13.95 based on data from 1993 to 1998.
Verification & Validation:	RSPA reviews/verifies data provided for accuracy and requests supplemental reports where shortcomings are indicated.
Comment:	RSPA is committed to reducing outside force damage to pipelines and will review progress to determine if we can "stretch" this goal further.

Hazardous Materials Incidents

Page 44

Measure:	Number of serious hazardous material incidents
Scope:	Serious reported hazardous materials incidents are defined by RSPA to be those that result in a fatality or major injury (for most purposes, an injury resulting in hospitalization) due to a hazardous material, closure of a major transportation artery or facility, or evacuation of six or more persons due to the presence of a hazardous material, or a vehicle accident or derailment resulting in the release of a hazardous material. This measure tracks only transportation related releases of hazardous materials that are in commerce. Volume of spills is not tracked, as this does not necessarily indicate risk.
Source:	Hazardous Materials carriers report data to RSPA for entry into the Hazardous Materials Information System (HMIS).
Limitations:	Data for all hazardous materials incidents is suspected of being incomplete due to under-reporting for minor incidents. Most reportable serious incidents are in the system, making this a more consistent measure for program management. However, it does not reflect all incidents.
Statistical Issues:	Although the number of incidents is likely to be underreported, such recording error is probably small in comparison to the annual variation due to chance. The annual variation in the number of failures from year to year due to chance has a regression standard error of 31.05 based on data from 1993 to 1996.
Verification & Validation:	RSPA verifies the data by periodic follow-up reviews of data entry by the manager of the Hazardous Materials Information System, and verification audits of the data entry process. RSPA crosswalks HMIS reports against the National Response Center log of accidents. RSPA is improving compliance with reporting requirements by correlating HMIS reports with FRA's Accident Report data and the HMIS telephonic data. RSPA plans to incorporate procedures to correlate HMIS reports with FHWA's Safetynet Accident File data.
Comment:	The RSPA rulemaking extending the jurisdiction of the HAZMAT regulations to include intrastate highway carriers, HM-200, will be fully implemented by FY 2000. The expected increase in the number of reported incidents will affect the performance measure, requiring a re-evaluation of the Department's goal.

Details on DOT Measures of Mobility

Highway pavement condition

Page 49

Measure:	Percentage of miles on the National Highway System (NHS) that meet pavement performance standards for acceptable ride.
Scope:	International Roughness Index (IRI) is compiled annually for every section of the NHS, using data reported from the States.
Source:	Data collected by the State Highway Agencies and reported to FHWA for the Highway Performance Monitoring System (HPMS). They are obtained from calibrated measurement devices that meet industry set standards. Measurement procedures are included in the HPMS Field Manual.
Limitations:	IRI data for the approved NHS exist from 1995 onward. Past data (1993 and 1994) contain some variation as this data was on the proposed, rather than the existing NHS. No NHS IRI data are available prior to 1993. The HPMS requires States to report IRI data every two years; however, following the requirements is not mandated, but voluntary.
Statistical Issues:	The major source of error in the percentages is probably the sampling error from selecting the segments of highway tested for smoothness. The annual variation in the percentage due to chance has a regression standard error of 0.270.
Verification & Validation:	FHWA validates the data based on consistency reviews. FHWA field offices perform annual reviews of the IRI process, including equipment and calibration checks.

Highway bridge condition

Page 51

Measure:	Percent of bridges on the NHS that are deficient
Scope:	Measure includes the number of deficient (structurally deficient and functionally obsolete) bridges on the NHS functional system divided by the total number of NHS bridges in the inventory, expressed as a percent.
Source:	Bridge information is collected by State DOTs and other bridge owners and provided to FHWA annually for inclusion in the FHWA maintained National Bridge Inventory (NBI). NBI contains data on 582,750 highway bridges.
Limitations:	Data are available from 1993 onward. NBI includes information on 582,750 bridges, including all 128,508 NHS bridges. It is the world's most comprehensive database of bridge information. States are required to update the system annually, but many States update quarterly. The system contains 95 data items for each of the bridges, and 20 of these items relate to bridge condition and appraisal. There are specific instructions as to how to assess bridges based on these items, including a grading scale from 0 to 9 with specific definitions and specific criteria to follow. This serves to reduce assessment subjectivity to a negligible level.
Statistical Issues:	Even with the item specific grading system, differences in the grading between individual inspectors and between inspection days are probably the largest component of variation in the percentages. Based on 1993-1998 data, the estimated regression standard error for year-to-year variation in the percentages due to chance is 0.677.
Verification & Validation:	DOT evaluates accuracy and reliability of the submitted NBI information through data checks and field reviews by both Headquarter and field office personnel. This is done as a part of FHWA's NBI, the National Bridge Inventory System (NBIS), and Highway Bridge Replacement and Rehabilitation Program. Evaluation of the State's compliance with the NBIS most often includes a sample of bridge inspection reports and a comparison of condition data with field visits to the bridge site. In addition, there is an edit update program that identifies potential data errors in the NBIS.

Highway congestion

Page 53

Measure:	Hours of delay/1000 vehicle miles traveled on Federal-aid highways.
----------	---

Scope:	Delay represents the difference between estimated actual travel speed and free flow travel speed that could be attained if there were no other traffic. Delay includes weekday and weekend travel combined. On other than freeways, delay also includes the delay due to traffic control devices – traffic lights and stop signs. Delay in hours per 1000 VMT is calculated on an individual section basis using the Highway Performance Monitoring System (HPMS) data and is summed to represent an annual average delay for all Federal-aid highways.
Source:	Data collected and provided by the State departments of transportation from existing State or local government databases, including those of Metropolitan Planning Organizations.
Limitations:	The delay calculation is modeled, based on traffic volume and capacity values such as number of through lanes, access control, and at-grade intersections. Minor lengths of Federal-aid highways on the lowest functional systems are not included in this analysis. Although nearly all States' data are included in the trend estimates, they do not include all States for all years. The estimate understates delay since it does not include delay due to incidents – crashes, etc. The estimate includes delay caused by traffic control devices since they reduce operating speed below what would otherwise be the free flow speed. Data exist for 1993 and later years.
Statistical Issues:	The primary source of uncertainty in estimating delay rates is the denominator. The numerator estimate of total delay, however, is also affected by modeling errors in the delay calculation. The annual variation in the delay rate attributable to random chance has a regression standard error of 0.332 based on data from 1993 to 1996.
Verification & Validation:	State-reported HPMS data are reviewed by FHWA for completeness, consistency, and adherence to reporting guidelines. Where necessary, and with close State cooperation, data may be adjusted to improve completeness, consistency, and uniformity.
Comment:	Nine hours of delay per 1,000 VMT could be equated to approximately 11 minutes of delay on 50 20-mile commuter trips in an urbanized area. In this example, the 11 minutes of delay easily could be the difference between the time it would take to travel the 20 miles at the posted speed with no stops and the time it would actually take during the height of the rush hour.

Intelligent Transportation Systems integration - (FY 2000 – FY 2001)

Page 55

Measure:	Number of Metropolitan Areas where integrated ITS infrastructure is deployed.
Scope:	<p>The level of integrated deployment in at least 75 of the nation's largest metropolitan areas will be established using a set of indicators that considered two factors: (1) How much ITS infrastructure is in place at each metropolitan area; and, (2) How much integration is going on at each area. The process for determining the level of "component" deployment in a metropolitan area will be made using a set of indicators that measure the magnitude of deployment for selected ITS components. Integration means sharing data between the different jurisdictions responsible for ITS infrastructure. Typically there are three: State DOTs responsible for management of freeways and incident management programs; city governments, which manage most of the traffic signal systems; and public transit authorities, which manage most bus and rail services. A metropolitan area is considered integrated if any two of the three major organizations employ technology to share and use transportation data to increase system capacity.</p> <p>Examples: 1) a city traffic signal system receiving data from the state freeway management center about the queues at freeway ramp meters and then adjusting the signal timings on the arterial streets, or 2) a transit agency providing the state freeway management center with the real-time location of the buses so that freeway speeds can be determined.</p>
Source:	Metropolitan ITS Deployment Tracking Database developed by the Oak Ridge National Laboratory for the ITS Joint Program Office. Data are collected by means of surveys from designated metropolitan areas.
Limitations:	This indicator is designed to track and encourage basic steps toward systems integration. However, it does not reflect the full breadth of integration activities. While it measures the existence of basic integration of essential components, it does not confirm that all possible or desirable integration links exist in a metropolitan area.
Statistical Issues:	These data come from sample surveys that, like all sample surveys, contain sampling and nonsampling errors.
Verification & Validation:	The DOT Joint Program Office reviews deployment tracking indicators and methodology. Results are distributed to DOT headquarters and field staff as well as to state and local survey responders for confirmation of accuracy and completeness before the final reports are issued. Survey construction and data collection

Comment:	procedures will be improved as a result of process feedback from each survey iteration. This indicator is meant to provide a basic, easy to understand gauge of integrated ITS deployment. This measure replaced the initial indicator of ITS integration (Increase in the level of ITS integration in 6 metropolitan areas above 1997 baseline) in FY 2000.
----------	---

Intelligent Transportation Systems integration - (FY 1999)

Page 55

Measure:	Increase in the level of Intelligent Transportation System (ITS) integration in 6 metropolitan areas above 1997 baseline.
Scope:	The process for determining the level of ITS integration in 6 of the nation's largest metropolitan areas will be made using a set of indicators that measure the magnitude of integration among selected ITS components. From a comprehensive perspective, a total of 34 integration links between nine ITS components can be identified. Only 11 of these links have been used to calculate an integration score for a metropolitan area. These 11 links address interactions among three agencies: State DOTs responsible for management of freeways and incident management programs; city governments, which manage most of the traffic signal systems; and public transit authorities, which manage most bus and rail services. Individual indicators are constructed for each of these 11 links. These individual indicators include such measures as the percentage of agencies that provide information to other agencies in an area for use in managing traffic, the percentage of agencies that participate in a freeway management program. The average of each of these 11 individual indicators is computed and used to measure the level of integration. The change over time is tracked by comparing the level of integration in an area for the 1997 base line.
Source:	Metropolitan ITS Deployment Tracking Database developed by the Oak Ridge National Laboratory for the ITS Joint Program Office. Data are collected by means of surveys from designated metropolitan areas.
Limitations:	This is an initial indicator designed to track and encourage basic steps toward systems integration. It does not reflect the full breadth of integration activities. The 1997 data displayed in the graph are based on actual counts from the deployment tracking surveys, with an overall response rate of 81%. However, the data would be relatively unaffected if the response rate were increased to 100%, because metropolitan areas that did not respond are generally those expected to have little or no ITS infrastructure deployed.
Statistical Issues:	This measure is an average of 11 indicators that may each have unequal variances. Also, the indicator might perform better if weights were applied to give more importance to more highly populated metropolitan areas (or other weights could be developed based on program goals.)
Verification & Validation:	The DOT Joint Program Office reviews deployment tracking indicators and methodology. Results are distributed to DOT headquarters and field staff as well as to state and local survey responders for confirmation of accuracy and completeness before the final reports are issued. Survey construction and data collection procedures will be improved as a result of process feedback from each survey iteration.
Comment:	This indicator served as an initial gauge of ITS integration. It was discontinued in FY 2000.

Runway pavement condition

Page 57

Measure:	Percent of runway pavement in good or fair condition (commercial service, reliever, and selected general aviation airports).
Scope:	Paved runways at the 3,300+ airports in FAA's National Plan of Integrated Airport Systems (NPIAS) are assessed for pavement condition. The NPIAS airports include all commercial service and reliever airports and those general aviation airports that are significant to national air transportation.

Source:	The FAA's Airport Safety Data Program (ASDP) provides extensive data about the facilities that are available at public-use airports. Data are provided approximately annually by FAA inspectors for airports certified under FAR 139. Data for other airports, including most public use general aviation airports, are provided under an FAA contract with the National Association of State Aviation Officials.
Limitations:	FAA contracts for a visual survey of the runways to categorize their condition based on criteria developed by the FAA Office of Airports. "Good" condition means all cracks and joints are sealed; "fair" condition means there is mild surface cracking, unsealed joints, and slab edge spalling; and "poor" condition means there are large open cracks, surface and edge spalling, and vegetation growing through cracks and joints. Since the reports are based on a visual inspection, underlying drainage or strength problems are not reported. However, these problems normally create surface defects that are visible. The more detailed PCI inspections require a section by section examination of the runway rather than an overall assessment used for this performance measure. FAA has been aggregating the ADSP data from all NPIAS airports only every several years for inclusion in the NPIAS report to Congress. This information exists for 1993, 1997, and 1998.
Statistical Issues:	The relatively subjective nature of judging pavement quality means this measure is subject to random variation due to measurement error.
Verification & Validation:	One-half of NPIAS airports have undertaken pavement condition index (PCI) surveys, which are more stringent and reliable (but still visual) than the surveys conducted under the Airport Safety Data Program. Comparisons of PCI and ASDP data show comparable results at the system level.
Comment:	Due to limited funds, the frequency of airport inspections by the National Association of State Aviation Officials (NASAO) has declined. Inspections by NASAO may cease during FY2000.

Aviation System Capacity

Page 59

Measure:	System capacity attributable to airport infrastructure at the 50 busiest airports.
Scope:	The estimated capacity of the runway systems of the 50 busiest airports, expressed as the number of annual aircraft operations that can be accommodated without incurring an excessive delay.
Source:	A computer model will be used by the FAA William J. Hughes Technical Center to calculate the capacity of the individual airports.
Limitations:	Only 14 airports will be simulated during 1999 and the capacity of the other airports will be estimated from tables, current planning documents, or other sources. The remaining airports will be simulated during 2000, 2001, and 2002.
Statistical Issues:	Since the computer simulation was not completed for the full sample of airports that was planned, there will be more sampling error in this measure than would have been the case if the entire sample was completed. This measure is also subject to errors in the specification of the model.
Verification & Validation:	FAA has extensive experience and full confidence in the computer simulation techniques. The accuracy of results depends on the validity of input assumptions and data, which are coordinated with the airport operators, airlines, and consultants.
Comment:	<p>Comment: preliminary information indicates that 50 airports may be more than optimum for monitoring the effectiveness of runway capacity enhancement programs because:</p> <ul style="list-style-type: none"> Runway congestion is a severe problem at a small number of airports. The busiest 25 airports account for almost all severe delays to air traffic. FAA is currently able to produce only 14 airports capacity analyses per year, so it will take 4 years to develop a benchmark for a 50-airport sample. Few uncongested airports plan to build new runways, so including them in the sample dilutes the perceived overall impact of new runways.

Aviation delays**Page 60**

Measure:	1. Volume-and FAA equipment-related delays per 100,000 flight operations. 2. Aviation delays per 100,000 activities.
Scope:	A FAA reported delay occurs when an aircraft is delayed fifteen minutes or more because of constraints that prevent the aircraft from making a scheduled landing. Delays are counted in five categories: FAA equipment, volume, weather, runway related, and other. The aviation delays measure used in FY 99 only included two categories – FAA equipment and volume. The measure used for FY 00 and FY 01 includes all five categories. Delays due to airline equipment are not considered. “Activities” are total facility activities, as defined in Aviation System Indicators 1997 Annual Report. Total facility activities are the sum of en route and terminal facility activities.
Source:	FAA air traffic facilities report the data to headquarters, which incorporates the data into the Air Traffic Operations Management System.
Statistical Issues:	There is no significant error in either the delay counts (numerator) or the flight operations data (denominator) for this rate. However, random variation in aviation delays results in a significant variation in the delay rate from year to year.
Limitations:	By collecting information on delays of fifteen minutes or more, FAA does not capture the aggregate amount of system delay, but only the most significant delays.
Verification & Validation:	Data is analyzed and checked by an Air Traffic Service headquarters office on a daily basis to ensure accuracy of the information reported.
Comment:	Total delays in all five categories are what the travelling public experience.

All Weather Access to Airports**Page 63**

Measure:	Number of published GPS landing approaches
Scope:	This performance measure counts the total number of published satellite-based landing approaches.
Source:	Internal FAA Aviation Systems Standards tracking system.
Limitations:	This is an output measure rather than an outcome measure. Individual use of GPS approach procedures is not tracked by current information systems. Although it may be impossible to measure the exact benefits to users, increased schedule reliability for commuters and air taxis, as well as improved access for general aviation will result from increasing the number of published approaches.
Statistical Issues:	There is no significant error in the counts of GPS landing approaches. However, random changes in the number of approaches caused by external factors results in random variation in the count from year to year.
Verification & Validation:	Productivity numbers are compared and validated monthly by FAA (Aviation Standards National Field Office and National Flight Data Center).

All Weather Access to Airports**Page 63**

Measure:	Number of runways that are accessible in low visibility conditions.
Scope:	This performance measure counts the total number of airport runways with published ground-based and/or satellite-based landing systems. The intent of this measure is to reflect increased accessibility using satellite-based technology.

Source:	Internal FAA Aviation Systems Standards tracking system.
Limitations:	Increasing the number of runways with satellite-based landing systems as well as augmenting existing satellite-based landing systems with vertical altitude guidance will improve access to airports and increase schedule reliability. Both improved access and increased reliability are considered benefits to the aviation industry and the individual air traveler. However, individual use of landing systems is not tracked by current FAA information systems. In addition, aircraft must be appropriately equipped to use the new technology. The FAA does not track these equipment additions.
Statistical Issues:	There is no significant error in the counts of published landing systems. However, like the above measure, random changes in the number of published approaches result in random variation in the count from year to year.
Verification & Validation:	The number of airport runways with a satellite-based landing system is computed monthly by Aviation Systems Standards.

Essential air service

Page 65

Measures:	1. Percent of subsidized communities with at least 2 round trips/day, 6 days/week (12 round trips/week) 2. Percent of subsidized communities with at least 3 round trips/day, 6 days/week (18 round trips/week).
Scope:	The measure shows the number of weekly round trips at subsidized EAS communities in the continental U.S. EAS communities are those that were on the certificated airline map in 1978.
Source:	Air carrier filings, community and carrier notification.
Limitations:	Service frequency is closely associated with program funding levels and the number of EAS communities that require subsidy; and the number of communities may change. Service frequency may also be affected by conditions such as an air carrier going out of business, airline strikes, or system shutdown. DOT's goal assumes a fairly constant level of communities in the base (76 in 1998). This measure will not show instances in which the Department is successfully able to effect a carrier transition to commercially viable service without a subsidy. Data has only been gathered for 1996 and later years.
Statistical Issues:	There is no significant error present in the subject data.
Verification & Validation:	Continued contact with civic parties, carrier officials, and Congressional staffs.
Comment:	Consideration of alternate strategies or performance measures may be prompted by the research study, <i>Economic Evaluation of the Impact of Air Service on Small Metropolitan and Rural Communities</i> , to be completed in 2000, or by other developments such as the state of aircraft manufacture.

Maritime navigation

Page 66

Measure:	Percentage of total operating days that marine aids to navigation are available for use on U.S. navigable waters.
Scope:	Includes all federally maintained aids to navigation except those on the Western U. S. Rivers. Privately placed and maintained aids to navigation are not included.

Source:	Coast Guard Aids to Navigation Information System (ATONIS)
Limitations:	This measure employs “discrepancy days” for individual aids. As such, it tends to overstate the discrepancy time of the aids to navigation system: a single aid outage usually does not degrade a waterway’s entire aid system and vessels are still able to transit normally. Complete system outages are rare, and usually result from severe weather incidents such as hurricanes. Western River system (roughly described by the Mississippi River system) aids to navigation are not included because of continuous changes in river conditions during the year, for example, river flow and silting patterns, which causes the need to continuously respond by changing aids positioning. Essentially, there is no position standards for these aids, against which to measure on-station position availability.
Statistical Issues:	As noted in the Limitations statement, this measure most likely creates a bias of overstating the discrepancy time of the aids to navigation system.
Verification & Validation:	Data entry software limits data entry errors. There is no structured or automated system for verification and validation. However, verification and validation is conducted at various levels – district managers review data periodically to identify missing and incomplete data. Program managers also review discrepancy data.
Comment:	This is an output measure, but it is closely related to the level of mobility afforded to vessels transiting navigable waterways.

Maritime navigation

Page 66

Measure:	Total number of navigational accidents - maritime collisions, allisions, and groundings.
Scope:	The measure includes collisions, allisions, and groundings of freight and tank ships over 500 gross tons. Intentional groundings are excluded. All U.S. and foreign ships in U.S. waters are included.
Source:	Coast Guard Marine Safety Information System (MSIS). Ship operators, crew, and pilots report accidents directly to the Coast Guard.
Limitations:	1999 data are preliminary estimates base on approximately 7 months of data. Data was extrapolated for a full year based on historical data patterns. It is probable that some events are not reported to the Coast Guard. This number is unknown. Serious events such as major collisions and grounding are more likely to be reported than minor events such as a temporary grounding where a vessel could remove itself without assistance. Duplicate event entries are sometimes entered into MSIS, and some events are mistakenly omitted or coded incorrectly. Verification procedures strive to correct these errors, but it is probable that a small number are not corrected. Because this count of incidents is not normalized to exposure, it does not provide a sensitive indicator of changes in risk.
Statistical Issues:	The primary source of uncertainty in estimating the performance measure is probably the response error. While there is a strong penalty for not reporting fatalities, the Coast Guard has initiated violation cases against 351 individuals and companies during the last 8 years for failure to report marine casualty cases. The number of unreported vessel casualties is greater than or equal to the number detected for violations. The measure, even with the changes suggested for FY 2001, can be subjective due to the judgement used by investigators to determine the primary nature of an accident. Differences between investigators will therefore be a major source of variation in the accident numbers.
Verification & Validation:	Verification and validation occurs at several levels. Edit checks within MSIS can detect some incorrect or missing data and force review and correction before data entry is completed. Selection lists for certain data fields also reduce the opportunity for data entry error. All investigations go through one level of review at the field unit for accuracy. Investigations of serious marine casualties are also usually reviewed at district and headquarters offices. The headquarters Data Administration staff conducts periodic quality control checks to identify entry errors such as missing data or miscoding, and corrects any errors identified. Each investigation involving a vessel accident is reviewed before it is included in the measure. Errors identified are referred to either the Data Administration staff or the Investigations and Analysis staff for correction.

Comment:	The target for collisions, allisions and groundings was originally based on data sorted according to the primary nature of the accident first reported to the Coast Guard. The Coast Guard has revised its analytical procedures to count for instances where collisions, allisions and groundings, not initially reported as the primary nature, but later confirmed to be a significant component in a chain of events, are included in the measure. For example, an incident initially reported as a pollution event, and subsequently confirmed to have been the result of a collision, allision, or grounding, would not have been counted under the old methodology, but is reported under the new methodology. This new methodology provides a more accurate picture of vessel traffic. The new target represents the same proportional reduction as our original target: 10% over 5 years.
----------	--

Impediments to port commerce**Page 68**

Measure:	1. Percentage of ports reporting landside impediments to the flow of commerce. (1999, 2000) 2. Percentage of ports reporting land- and waterside impediments to the flow of commerce. (2001)
Scope:	<p>The top 58 U.S. deep draft ports (including the top 50 U.S. ports, the top 25 container ports, and the 14 DOD-designated strategic ports). Some ports may be in more than one category. Also includes the top 23 shallow draft ports (based on tonnage). In total, 81 deep and shallow draft ports are captured in the measure. Examples of landside impediments include insufficient truck turning radii, inadequate signage, inadequate bridges, etc. Waterside impediments include channel depth, terminal depth, and dredging requirements.</p> <p>The impediments to port commerce measure used for FY 1999, and scheduled for FY 2000, examine only landside impediments at the 58 deep draft ports. The measure for FY 2001, which includes land- and waterside impediments at the 81 deep and shallow draft ports, is currently under development. A new survey instrument is being developed to collect the necessary information from a larger and more diverse sample of ports.</p>
Source:	A FY 1999 analysis by MARAD. Data was derived from Metropolitan Planning Organizations; State DOTs' transportation improvement plans submitted under ISTEA/TEA-21; the FHWA Connectors Performance and Investment study; and, Border and Corridor projects awarded under TEA-21; and site visits by MARAD personnel.
Limitations:	Ports may be unwilling to acknowledge impediments because it could damage their competitive position relative to other ports. The wide variety of impediments and number of intermodal projects make comprehensive annual reporting prohibitively expensive. Terminals within a specific port area may have widely varying impediments. There may be limitations on the availability of data on private-sector lead projects that have no federal involvement or funding.
Statistical Issues:	The performance measure for FY 1999 and 2000 is calculated using data from a variety of sources, since US ports are not required to provide MARAD with data on impediments. As such, the data probably exhibit gaps in information, duplication, and inconsistency in definition. The proposed survey to gather the data for FY 2000 is designed to gather more consistent information.
Verification & Validation:	Data sources listed above are mainly from public reporting requirements to the Federal Government. The accuracy and completeness of the data can be verified through agency sources.

St. Lawrence Seaway lock availability**Page 70**

Measure:	Percentage of days in the shipping season that locks are available.
Scope:	Includes "downtime" (delay or prohibition of transiting) for transit of the U.S. sectors of the St. Lawrence River throughout the navigation season (late March to late December). Downtime is measured in minutes/hours of delay for weather (visibility, fog, snow, and ice); vessel incidents (human error, electrical and/or mechanical failure); water level and rate of flow regulation; lock equipment malfunction.
Source:	SLSDC gathers the data from Lock Operations Records.
Limitations:	SLSDC reports data directly from observation. The indicator does not cover all characteristics of performance of the St. Lawrence Seaway.

Statistical Issues:	It should be noted that “downtime” for transit encompasses the two U.S. seaway locks in Massena, N.Y. The SLSDC’s lock operation unit gathers primary data for all vessel transits through these two U.S. locks.
Verification & Validation:	SLSDC verifies and validates the accuracy of the data through review of 24 hour vessel traffic control computer records, radio communication between the two Seaway entities and vessel operators; and video and audio tapes of vessel incidents.
Comment:	SLSDC influences the measure primarily through capital planning, and consistent facilities maintenance and investment.

Amtrak ridership – (FY 2000 - FY 2001)

Page 72

Measure:	Number of passengers on Amtrak’s intercity routes
Scope:	The measure includes all revenue paying passengers on intercity routes.
Source:	Amtrak Annual Report.
Limitations:	Data collection relies on accuracy of Amtrak report. Ridership is an outcome indicator that reflects a variety of factors, not insignificantly the capital investment of the federal government. Operational decisions of Amtrak and the availability and cost of alternative modes of transportation also influence ridership.
Statistical Issues:	Chance variation from year to year, as estimated by the regression standard error from 1993-1998, is 0.77. This calculation assumes stable operations over the five-year period; since new runs and lines are added and removed fairly often, the standard error is only a rough approximation.
Verification & Validation:	Amtrak conducts monthly verification and validation of data.
Comment:	A 3.6 million increase in ridership is projected from 1998-2001 as a result of the initiation of the Northeast Corridor high-speed rail service.

Amtrak ridership – (FY 1999)

Page 72

Measure:	Amtrak trip time between NY and Boston.
Scope:	Scheduled trip time between New York City and Boston.
Source:	Published timetables.
Limitations:	The Acela trainset is currently undergoing safety qualification tests.
Statistical Issues:	The published trip time is likely to slightly underestimate the true average; delays are likely to occur more frequently and cover a wider time span than early arrivals. But such effects are, for commercial reasons, likely to be small.
Verification & Validation:	Published timetables.
Comment:	Amtrak introduced limited electrification service in January 2000. These trains reduced scheduled trip time to four hours. Acela high-speed service is expected by mid-year, reducing trip time by at least an additional 45 minutes. The three-hour goal will be reached when three major track reconfigurations are completed at New Rochelle, NY and Stamford and New Haven, CT.

Amtrak ridership – (FY 1999)

Page 72

Measure:	Percentage of Amtrak trips arriving on time.
Scope:	System-wide on-time arrivals.

Source:	Amtrak Annual Report.
Limitations:	Data collection relies on accuracy of Amtrak report.
Statistical Issues:	Annual variation due to chance in this percentage, as estimated by regression standard error from 1993-1998, is 2.62.
Verification & Validation:	Amtrak conducts monthly verification and validation of data.
Comment:	A nine-percentage point increase in on-time performance is projected from 1998-2001, due to the initiation of the Northeast Corridor high-speed rail service. Performance increases are also attributable to the remaining Amtrak operation as freight congestion declines, as a result of the recent rail mergers.

Amtrak ridership – (FY 1999)

Page 72

Measure:	Customer satisfaction index.
Scope:	The customer satisfaction index (CSI) includes on-time performance, comfort, ride quality, and equipment condition.
Source:	Amtrak Annual Report. The CSI is prepared by Amtrak customer surveys. The surveys allow customers to report on a variety of service indicators
Limitations:	Data collection relies on the accuracy of the Amtrak Report.
Statistical Issues:	Surveys and indices pose difficult statistical problems; there are many potential sources of uncertainty. In this case, the annual random variation, as estimated by the regression standard error from 1993-1998, is 0.59.
Verification & Validation:	Amtrak conducts monthly verification and validation of data.
Comment:	A five-percentage point increase in CSI is projected from 1998-2001, due to the initiation of the Northeast Corridor high-speed rail service. Improvements in technology associated with high-speed service are also expected to improve service.

Amtrak ridership – (FY 1999)

Page 72

Measure:	Number of intercity and commuter trains scheduled along the most congested segments of the Washington/Boston corridor.
Scope:	Intercity and commuter trains scheduled along the most congested segments of the Washington-Boston corridor.
Source:	Published timetables.
Limitations:	None.
Statistical Issues:	None.
Verification & Validation:	Published timetables.
Comment:	Projected increase is 365 trains per day by 2005.

Transit ridership – (FY 1999)

Page 75

Measure:	Revenue vehicle hours of service (rail and non-rail, in millions).
Scope:	Includes revenue vehicle hours of service on publicly funded bus, transit rail, commuter rail, ferry, paratransit, and vanpools in urbanized areas.

Source:	National Transit Database (NTD). By law, each year, about 600 public transit operators report audited data for urbanized areas.
Limitations:	Service hour data is recorded daily. Recording problems can arise during vehicle breakdowns and service interruptions.
Statistical Issues:	The primary source of uncertainty in this estimate derives from the accuracy of the audited data. This uncertainty is almost surely small, but difficult to estimate without the special study.
Verification & Validation:	An independent auditor and the transit agency's CEO certify that hours data reported to the NTD are accurate. Analysts validate data. Hours data are compared to hours in previous years and are cross-checked with other related operating and financial data in the report. Validation problems must be corrected or explained by the reporter. Vehicle hours are reviewed as part of FTA's Triennial Review and Vehicle Fleet Management Requirements.
Comment:	None.

Transit ridership – (FY 2000 – FY 2001)

Page 75

Measure:	Billion transit passenger miles traveled.
Scope:	Includes revenue-passenger miles on publicly sponsored bus, transit rail, commuter rail, ferry, and vanpools in urbanized areas.
Source:	National Transit Database (NTD), with information gathered from transit operators.
Limitations:	Data is self-reported by transit agencies using an FTA-approved sampling methodology. Although most data is reported in the National Transit Database each year, sample cycles may be annual, every three years, or every five years depending on the size of the urban area and the number of vehicles operated. Ridership is an outcome indicator that reflects a variety of factors, including the capital investment of the federal government. Ridership is also influenced by operational decisions of transit authorities, and the availability and cost of alternative modes of transportation.
Statistical Issues:	The sources of uncertainty include sampling error, annual chance variation, and auditing issues. The regression standard error from 1993-1998 indicates that the magnitude of the combination of the first two sources of error is about 0.41.
Verification & Validation:	An independent auditor and the transit agency's CEO certify that data reported to the NTD are accurate. FTA also compares data to key indicators such as vehicle revenue miles, number of buses in service during peak periods, etc.
Comment:	None.

Bus and rail transit fleet condition

Page 77

Measure:	1. Average condition of motor bus fleet. 2. Average condition of rail vehicle fleet.
Scope:	The measure includes bus, demand response, and rail fleets
Source:	National Transit Database (NTD), with information gathered from transit operators; Transit Economic Requirements Model (TERM), which estimates average vehicle condition using NTD data.
Limitations:	Average vehicle condition may not fully reflect the average condition that transit passengers face, since vehicles in worse condition tend to be utilized less. There are also lags in reporting of data to the NTD (thereby requiring preliminary estimates for recent years) and in the effects of federal government capital assistance (since it may take five years from the time that funds are appropriated to the time that new or rehabilitated vehicles are placed in service)

Statistical Issues:	Condition is generated from NTD data using an econometric model, which in turn is based on a random national sample of vehicles. Average condition changes very slowly due to the steady replacement of vehicles and the relationships in the estimated model.
Verification & Validation:	An independent auditor and transit agency's CEO certify that data reported to the NTD are accurate. Data are also compared with fleet data reported in previous years, and cross-checked with other related operating/financial data in the report. The econometric model used to translate NTD data into average condition ratings is based on visual inspections of a national sample of bus and rail vehicles. The sample will need to be repeated periodically in the future in order to keep the econometric model current with developments in vehicle conditions.
Comment:	The use of average condition rather than average age is an improvement that is new to this Performance Report.

Transportation accessibility

Page 79

Measure:	1. Percentage of bus fleets that are Americans with Disabilities Act (ADA) compliant. 2. Percentage of key rail stations that are Americans with Disabilities Act (ADA) compliant.
Scope:	Accessibility for bus fleet means that vehicles are lift or wheel chair ramp equipped. Accessibility for key rail facilities is determined by standards for ADA compliance.
Source:	Data on bus accessibility is collected in the National Transit Database (NTD), with information gathered from transit operators. Data on rail accessibility is reported to FTA by the transit authorities.
Limitations:	Measure does not capture ADA compliance (or transportation accessibility) for modes other than transit.
Statistical Issues:	None.
Verification & Validation:	For bus accessibility, an independent auditor and the transit agency's CEO certify that data reported to the NTD are accurate. Data are also compared with fleet data reported in previous years, and cross-checked with other related operating/financial data in the report. Fleet inventory is reviewed as a part of FTA's Triennial Review, and a visual inspection is made at that time. FTA's Office of Civil Rights conducts oversight reviews in order to verify the information on key rail station accessibility which has been self-reported by the transit authorities.
Comment:	FTA will primarily influence the goal through Federal transit infrastructure investment, which speeds the rate at which transit operators can transition to ADA compliant facilities and equipment.

Details on DOT Measures of Economic Growth & Trade

Appalachian highway system

Page 85

Measure:	Miles of the Appalachian Development Highway System (ADHS) completed.
Scope:	Measure includes actual miles completed on the 3,025 mile ADHS, within 13 member States.
Source:	States submit annual status updates on ADHS miles completed within their State to the Appalachian Regional Commission (ARC). The ARC compiles the data.
Limitations:	This is an output measure.
Statistical Issues:	None.
Verification & Validation:	Completed by ARC.
Comment:	ARC estimates that the TEA-21 funding level will result in completion of approximately 37 additional miles each FY 1999 through 2003.

Flight route flexibility**Page 86**

Measure:	Percentage of flights that aircraft are able to fly off ATC-preferred routes.
Scope:	Published air traffic (ATC) preferred routes restrict many of the more heavily traveled routes between major metropolitan area in the National Airspace System. The removal of ATC preferred routes will positively impact the measure.
Source:	FAA Enhanced Traffic Management System.
Limitations:	The action of not assigning or eliminating preferred routes does not automatically make a contribution to the goal of aviation efficiency. It does provide flexibility to the industry and the potential for improved efficiency in certain situations.
Statistical Issues:	There is no significant error in the counts of ATC preferred routes. However, random changes in the number of ATC preferred routes, as well as changes in the number and distribution of airline flights, results in random variation in the measure from year to year.
Verification & Validation:	Air Traffic Service analyzes data collected by Volpe National Transportation Systems Center for air traffic facilities.

International air service**Page 88**

Measure:	Number of passengers (in millions) in international markets with open aviation agreements.
Scope:	These data are collected by DOT for all flight segments to/from a U.S. point. The data for this measure include all passengers on U.S. and foreign carrier flights to and from 36 "open-skies" countries and Canada. This indicator reflects (barring significant, unrelated macroeconomic and political influences) the extent to which the competitive environment promoted by DOT increases travel opportunities.
Source:	U.S. air carriers file domestic and foreign data in the T-100 system. Foreign carrier data are from the T-100F database. Foreign air carriers file data for all nonstop flight segments involving a U.S. point.
Limitations:	These data are considered a reliable measure of airline passenger traffic between the U.S. and foreign nations. The annual increase in air traffic, however, is affected by economic strength as well as market liberalization in bilateral aviation trade agreements. Furthermore, only part of the growth rate in open skies markets can be attributed to new traffic – some of the increase may reflect diversion of traffic from less competitive routes, with higher taxes and/or inferior service options. The goal of 3% annual growth reflects aviation analysts' judgment of the net impact of these agreements above the estimated growth expected in the industry. For these reasons, this measure must be considered more of a forecast than a "target" and program effectiveness will be assessed in greater detail both in the narrative of the annual performance report and in program evaluations (one is slated in 2000).
Statistical Issues:	Like other counts of aviation-related activities, there are no significant sources of systematic error in these data that have been quantified. However, random variation in the number and distribution of airline passengers, as well as the changes in the number of "open-skies" agreements, results in variation in the measure over time. The regression standard error in this variation for 1993 through 1998 is 0.78.
Verification & Validation:	Airlines are required to certify that these data are accurate. Also, these data are a 100% enumeration of traffic and capacity and can be verified for reasonableness against other databases, such as flight schedules.
Comment:	U.S. policy has favored the linking of networks. Networks allow improved service and marketing in many thousands of small city-pair markets. All of this traffic flows over flights captured by the T-100 and T-100F reports for international flights.

Great Lakes winter navigation**Page 89**

Measure:	Days critical waterways are closed due to ice.
-----------------	---

Scope:	7 waterways are designated critical to icebreaking on the Great Lakes based on historical ice conditions, volume of traffic, and potential for flooding due to ice dams on rivers. Winter conditions are defined by a standard severity index (-6.2 or milder defines average severity; more than -6.2 defines severe). The measure is the number of days critical waterways are closed for more than 24 hours due to ice.
Source:	Data comes from U.S. Coast Guard and U.S. Army Corps of Engineers observations. Waterways closure data is reported to the Ninth Coast Guard District by operating units via operational situation reports.
Limitations:	The data set associated with this measure is relatively small and simple; hence it is estimated to be fairly accurate. However, it is possible that small errors exist. This measure captures only Great Lakes winter navigation, and not all domestic icebreaking. The observation of closures in critical waterways is a surrogate for mobility over the whole Great Lakes waterway system.
Statistical Issues:	This particular performance measure is highly sensitive to the severity of winter weather, which will dramatically affect the ability to predict the number of days the waterways are closed due to ice. The term "waterway closure" is also subject to differences in definition by districts or sub-units reporting the data.
Verification & Validation:	Coast Guard district program managers review and validate data from situation reports and provide Coast Guard headquarters with an End of Season report.
Comment:	Great Lakes data reflect initial measurement methodology. Further refinements are being explored that will make this measure a more comprehensive gauge of winter navigation.

Commercial shipbuilding

Page 90

Measure:	Gross tonnage (in thousands) of commercial vessels under construction in U.S. shipyards.
Scope:	Includes all commercial vessels over 100 GT built in the United States.
Source:	Lloyd's Ship Particulars File; MARAD surveys of U.S. shipbuilding companies that construct vessels of 100 GT and over; Maritime Guaranteed Loan (Title XI) applications.
Limitations:	MARAD does not have a complete time series of the commercial orderbook. (Historical orderbook data available to MARAD only include vessels 1,000 GT and above; this excludes smaller commercial vessels built in the United States.) Beginning in 1997, the Lloyd's Particulars File became MARAD's primary source of data on the U.S. commercial orderbook; however, Lloyd's gross tonnage data are only available for vessels 100 GT and above and do not include certain types of vessels such as offshore boats, barges, drill rigs, etc. MARAD has also found inaccuracies of up to 60 percent in the Lloyd's data. Beginning in 1998, MARAD commenced conducting semi-annual shipyard surveys. While the overall response rate is approximately 40 percent, MARAD believes that the survey data from the 75 commercial shipbuilders provides an accurate representation of the status of the shipbuilding industry and related construction activity for vessels over 1,000 GT. (MARAD expects that new information-gathering techniques commencing in CY 1999 will obviate the low-response problem.) The data gaps are filled through extensive use of trade publications (e.g., <i>Lloyd's</i> , <i>Fairplay</i> , etc.)
Statistical Issues:	There is a disconnect between the vessels supposedly included in the performance measure (vessels over 100 GT) and the primary data source (vessels > 1,000 GT). There is a gap in the data for those vessels between 100 and 1,000 GT, which the Lloyd's Particulars File attempts to fill (but only for a subset for these vessels).
Verification & Validation:	Semi-annual surveys will be enhanced to address the data deficiencies described above by increasing the number of shipyards surveyed and by confirming information published in the trade press with shipyards. MARAD follows up shipyard surveys by telephone and other electronic means to clarify and/or verify certain responses to our surveys. Additionally, MARAD began to conduct on-site verification of the data in FY 1999.

Access to jobs

Page 92

Measure:	Number of employment sites that are made accessible by Job Access and Reverse Commute transportation services.
----------	--

Scope:	This measure assesses one part of the Job Access and Reverse Commute program – the number of employment sites made accessible that were not previously accessible. An employment site is considered accessible if located within ½ mile of services provided by the grantee. Employment sites must offer jobs that require a high school diploma or less. Services that make an employment site accessible may include, but are not limited to, carpools, vanpools, and demand-responsive services as well as traditional bus and rail public transit. The measure cannot account for those Job Access and Reverse Commute activities that encourage riders to use already existing sources of public transit. See comments.
Source:	Data are provided to FTA by grantees of the Job Access and Reverse Commute program on a regular basis.
Limitations:	The goal and measurement is a preliminary effort at capturing results of the Job Access and Reverse Commute program. Three elements are key to job access – the residence of the employee, the commute, and the job location. This measure includes the “goal” of the commute and the job, but it does not include the “starting line” of the commute, the rider’s home. Although jobs may be made more accessible to transportation services, these services may not provide access to potential workers’ communities. This measure also cannot account for improved accessibility due to lower fares or shorter commute times – it only addresses the gap in service delivery. FTA requires a greater level of precision from larger, urban grantees than rural grantees that may have fewer resources at their disposal.
Statistical Issues:	There are major problems in obtaining accurate estimates of the number of entry-level jobs within a half-mile of grantee services. Surveys are costly and prone to systematic biases. The uncertainty in this estimate is both large and difficult to quantify.
Verification & Validation:	FTA approves the local methodologies for collecting this data, ensuring that the data is consistent and comparable.
Comment:	This goal and its measurement will have to be revisited and revised as the Department gains more experience with the program. Job access programs are difficult to measure, as there are very different obstacles to overcome. Services can make employment sites accessible by closing a spatial gap, reaching new geographic locations. They can also close a temporal gap, providing services at a time when other services are not offered. Because this program works with nontraditional projects and grantees, it can include a great variety of transportation services.

Transportation and education

Page 94

Measure:	Number of students graduating with transportation-related advanced degrees from universities receiving DOT funding.
Scope:	University Transportation Center (UTC) data includes recipients of Masters and Ph.D. degrees in programs considered to be transportation related.
Source:	UTC data to be derived from university records provided to RSPA as part of the UTCs’ grant application.
Statistical Issues:	There is a possibility of undercounting, due to difficulty in specifying degree programs that are transportation-related. Additionally, some universities may not fully comply.
Limitations:	No data currently exists for the UTC program or for other education programs that can result in graduate degrees.
Verification & Validation:	Comparison with data reported for all degree programs by host universities and specific reports on each recipient of an advanced degree.

Transportation and Education

Page 94

Measure:	Cumulative number of students (in thousands) reached through the Garrett A. Morgan Technology and Transportation Futures Program.
Scope:	Includes students of all ages reached through specific activities such as internships, job shadowing, career days, video conferences, classroom visits, and teacher externship visits that inform them of the opportunities available in the transportation field and ensure that they have the skills and knowledge required for transportation jobs.

Source:	RSPA maintained database to aggregate responses from program organizers.
Limitations:	The inherent nature of this measure does not allow us to gauge the quality of contacts made with students "reached" or provide a means to track outcomes in terms of students entering the transportation field as a direct result of the activities sponsored through the Garrett A. Morgan Technology and Transportation Futures Program.
Statistical Issues:	Some variability is inevitable in classroom attendance counts, videoconferences, and other measures of exposure. But this uncertainty should be small.
Verification & Validation:	RSPA works to ensure that the quantitative data being reported is complete and accurately reflects the associated student activity before it is entered into RSPA's database.
Comment:	We are working to develop additional ways of reaching students and improving on this overall measure.

Disadvantaged and women-owned business contracting

Page 96

Measure:	<ol style="list-style-type: none"> 1. Percent share of the total dollar value of DOT direct contracts that are awarded to women-owned businesses. 2. Percent share of the total dollar value of DOT direct contracts that are awarded to small disadvantaged business.
Scope:	Includes contracts awarded by DOT through direct procurement (i.e., not including contracts issued by grantees)
Source:	Data from the Contract Information System (CIS) as reported by all DOT contracting activities to the Federal Procurement Data Center (FPDC).
Limitations:	Contracting data is reported by procurement office directly into the CIS.
Statistical Issues:	There is no significant error present in the subject data. However, random variation in the number of DOT contracts as well as the number of women-owned and small-disadvantaged businesses each year results in some random variation in these measures from year to year. The regression standard error for 1993-1998 is 0.73 for women-owned businesses and 3.07 for small-disadvantaged businesses.
Verification & Validation:	SBA conducts verification and validation of data by comparing annual reports submitted by DOT against FPDC data.
Comment:	As a result of the Adarand Supreme Court case, the direct contracting and DOT regulations affecting affirmative action programs have been revised. This may significantly alter small disadvantaged and women owned businesses ability to participate in federal contracting. The changes taking place in FY 1999 and beyond may reduce the participation levels and the goal may have to be adjusted.

Details on DOT Measures of Human & Natural Environment

Mobile source emissions

Page 101

Measure:	Tons (in millions) of mobile source emissions from on-road motor vehicles.
Scope:	Figure is the sum of on-road mobile source emissions of carbon monoxide, hydrocarbons, nitrogen oxides, and particulate matter less than 10 microns in diameter (PM-10).
Source:	National Air Quality and Emissions Trends Report published annually by EPA. (EPA uses data from FHWA's Highway Performance Monitoring System - HPMS.)
Limitations:	On-road mobile source emissions estimates are modeled using vehicle data. Past data contain some variations due to changes in methodology used to obtain these on-road mobile source emissions estimates. EPA revises emission estimates periodically based on revised methodology. In 1999, EPA increased the annual emission burden trend based on the knowledge that heavy duty diesel trucks manufactured since the early 1990's produce higher emissions during high speed operations. Emissions data are reported in a 2-year time lag. Indicator captures all major mobile source emissions from on-road vehicles. It does not capture off-road mobile sources, such as agriculture and construction machinery, lawn mowers, aircraft, trains, and boats.
Statistical Issues:	The EPA's use of a mathematical model poses issues of model validation. The annual variation in the model's estimates, as measured by the regression standard error, is 2.57. The HPMS data used as input to the model are subject to sampling and nonsampling errors.
Verification & Validation:	EPA conducts verification and validation of data. FHWA field offices perform annual reviews of HPMS data that EPA uses as a part of its model.
Comment:	The National Ambient Air Quality Standards (NAAQS), as revised in July 1997, may create new challenges for DOT in meeting the air quality goal. Targets may need to be modified to reflect these changes.

Greenhouse gas emissions

Page 103

Measure:	Metric tons (in millions) of carbon equivalent emissions from transportation sources.
Scope:	Measure includes GHGs that will be subject to the Kyoto Protocol, if ratified by the Senate (e.g., CO ₂ , CH ₄), but not other GHGs (e.g., water vapor). Includes emissions from international travel and shipping to and from the U.S., but not from that between other countries. Does not include emissions from non-transportation mobile sources such as farm and construction equipment.
Source:	<i>Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-1997</i> , published by EPA, supplemented with <i>Emissions of Greenhouse Gases in the United States 1998</i> , published by DOE's Energy Information Administration (1998 data). Estimates are based on data from EPA, DOE, and other agencies.
Limitations:	GHG emissions are estimated based on DOE estimates of aggregate supply of energy products such as motor gasoline and distillate fuel oil. Further disaggregation (e.g., of transportation modes and other uses such as agriculture) is not always available. Related "upstream" emissions and sequestration (e.g., from petroleum refining) are in separate categories. Procedures for calculating and applying GHG credits and permits have not yet been established.
Statistical Issues:	These data are external to DOT. They are subject to both sampling and nonsampling errors.
Verification & Validation:	EPA and DOE conduct verification and validation of data. DOT will participate as appropriate in reviewing data, methodology, and results.

Comment:	If entered into force, the Kyoto Protocol ("the Protocol") to the United Nations Framework Convention on Climate Change (UNFCCC) would establish a binding limit on aggregate U.S. emissions of six GHGs during 2008-2012, but would not establish any sector-specific limits. However, the Protocol would defer to the International Maritime Organization (IMO) and the International Civil Aviation Organization (ICAO) for the development of guidelines for reducing emissions associated with combustion of marine and aviation bunker fuels, respectively. The Protocol would provide for the transfer of emissions credits and/or permits between sectors and countries, but would require further development of accounting and other procedures. Pending the establishment of a national goal for greenhouse emission, DOT will monitor this measure only.
----------	--

Energy Efficiency

Page 105

Measure:	Transportation energy consumption (in quadrillion BTU's) per trillion dollars of Real Gross Domestic Product (GDP).
Scope:	Measure includes primary consumption of petroleum for transportation, in quadrillion BTUs. This does not account for petroleum-produced electricity that is used in transportation; however, this is less than 1% of petroleum use. Petroleum use is normalized to real GDP, in constant 1992 dollars.
Source:	U.S. Department of Energy <i>Annual Energy Review 1998</i> .
Limitations:	Energy consumption does not include petroleum-produced transportation electricity. Measure does not capture the fraction of this petroleum use that is imported, nor does it capture actual energy efficiency (BTUs per passenger-mile-traveled).
Statistical Issues:	These data are external to DOT. They are subject to both sampling and nonsampling errors.
Verification & Validation:	Data is taken from external sources, which conduct their own verification and validation.
Comment:	Petroleum use is normalized to the nation's real GDP in order to capture the nation's economic exposure to petroleum use in transportation. When transportation petroleum use is normalized to passenger-miles-traveled, the trend is upwards – energy use per passenger is becoming less efficient. No goal has been set for transportation energy use per real GDP at this time, pending further study and consultation with DOE.

Wetland protection and recovery

Page 107

Measure:	On a program-wide basis, acres of wetlands replaced for every acre affected by Federal-aid Highway projects (where impacts are unavoidable).
Scope:	Measure includes wetlands associated with all Federal-aid highway projects each fiscal year. To be included, wetland replacement (or investment in a wetland bank) must have begun.
Source:	State DOTs input Federal-aid related wetland degradation and replacement data into either a locally developed wetland mitigation databases or the FHWA Wetlands Management Database. FHWA compiles the final data.
Limitations:	Data only exists on Federal-aid related wetland replacement. Also, uniformity of the data is not guaranteed, as it is subject to interpretation by the reporting State DOTs. In particular, there is no uniform understanding of what should be reported as mitigation acreage. The FHWA has provided guidance on mitigation activities to report and will soon issue the Wetlands Management Database that should reduce the current variations in data received from the States. Data on wetland replacement is available for the past four fiscal years (FY 1996 - FY 1999).
Statistical Issues:	The non-uniformity of the data is problematic. Definitional ambiguity also makes formal statements of statistical uncertainty problematic.
Verification & Validation:	Data are gathered from established mitigation amounts required by section 404 permits that states must acquire for their projects. In addition, FHWA provides guidance to help states consistently report mitigation data. This process will be further improved through a standard mitigation database under development for the states. At present, there is no external audit of state data.

Comment:	All Federal agencies (including DOT, FHWA, and other modes) must comply with National Environmental Policy Act (NEPA) and the Clean Water Act (specifically section 404(b)(1) of the CWA) regarding disruption of wetlands. These laws require agencies to identify project alternatives that would avoid or minimize impacts to wetlands as a first consideration. These alternatives are subjected to analysis under both NEPA and the Clean Water Act. Under the law, these alternatives must be chosen unless the project sponsors clearly demonstrate that they are not viable because they do not meet the project purpose and need or will lead to other more significant environmental impacts. If, in compliance with the law, wetland disruption is unavoidable, FHWA then works to achieve this goal of wetland replacement.
----------	---

Livable communities - transit service

Page 109

Measure:	Percent of urban population living within a quarter mile of a transit stop with service frequency of 15 minutes or less (non-rush-hour).
Scope:	A transit stop is defined as a bus stop, but does not include rail stations unless associated with a bus stop.
Source:	FTA compiled information from bus schedules across the country. Population statistics come from the Census Bureau. Information from both of these sources was formatted using the Geographic Information System.
Limitations:	Transit stops do not include rail stations (such as light rail or subway). However, rail stations are almost always served by bus lines, so most persons who live near a rail station also live near a bus line.
Statistical Issues:	The extrapolation of population statistics from the Census Bureau at a level fine enough to support inferences within a geographic radius of a quarter mile is difficult. The measurement aspects of this estimate require careful examination.
Verification & Validation:	Under development.
Comment:	This measure is an attempt to gauge the use of transit for basic mobility. The Federal Transit Administration is working to develop the Transit Performance Monitoring System. Fully instituted, the TPMS will allow the Department to measure not only how many people live close to public transit, but also how many people use public transit for basic mobility. DOT is assessing the validity of this measure.

Aircraft noise exposure

Page 111

Measure:	Number of people in the U.S. (in thousands) who are exposed to significant noise levels (65 decibels or more).
Scope:	Residential population exposed to aircraft noise above Day-Night Average Sound Level of 65 decibels around the 250 U.S. airports with the greatest number of commercial jet take-offs and landings.
Source:	A statistical modeling technique (Nationwide Airport Noise Impact Model, or NANIM) is applied using the 250 largest civil airports with jet operations in the U.S. Flight activity forecasts, commercial fleet mix and population projections are developed from the Terminal Area Forecast (TAF) and Metropolitan Statistical Area (MSA) population forecasts. 1990 census data are subjected to multiple source updates as part of an international study application. FAA's Part 91 database supplies the number of hushkitted and re-engined Stage 2 aircraft. Noise contour information is derived from the FAA Integrated Noise Model (INM) and generic procedures used in the FAA Area Equivalent Method (AEM).
Limitations:	No actual count is made of the number of people exposed to aircraft noise. No military or general aviation aircraft are included in the FAA's model.
Statistical Issues:	This measure is derived from model estimates that are subject to errors in model specification. Additionally, the model does not estimate total population exposure, but instead estimates change over time since a benchmark established by the EPA in the mid 1970s. A new model is under development that is expected to give higher estimates than the current model.
Verification & Validation:	The Integrated Noise Model has been validated with actual acoustic measurements at both airports and other environments such as areas under aircraft at altitude. External forecast data are from primary sources.

Maritime Oil Spills

Page 113

Measure:	Gallons spilled per million gallons shipped, by maritime sources.
Scope:	Spills from regulated vessels and waterfront facilities are counted; other spills are not. Oil spills of 1 million gallons or more are excluded from data since they are rare (they do not occur every year) and would have an inordinate influence on statistical trends. The 1 million gallon threshold is the same as that used in the National Contingency Plan for defining major oil spills in coastal waters.
Source:	Spill amounts (numerator) is obtained from the Coast Guard Marine Safety Information System (MSIS). Spills are initially reported to the Coast Guard National Response Center by the spiller or, in some cases, by third parties. Data on waterborne oil shipments (denominator) is from US Army Corps of Engineers "Waterborne Commerce Statistics".
Limitations:	1999 data are preliminary estimates based on approximately 7 months of data. Data was extrapolated for a full year based on historical data patterns. It is probable that some spills are not reported. Large spills that impact a large area and number of people are more likely to be reported than small spills. The amount of oil spilled may be over or under-reported since it is difficult to precisely estimate. Spillers may also tend to underestimate the amount of oil spilled. Duplicate spill entries are sometimes entered into MSIS, and some spills are mistakenly omitted or entered incorrectly. Verification procedures strive to correct these errors, but it is probable that a number are not corrected. The error rate for volume of oil spilled is estimated to be less than 5% because large spills receive a high level of review and account for most of the volume spilled. By excluding non-regulated sources and major oil spills, the measure does not capture the entire overall amount spilled annually. The measure is more meaningful for program management, but may under-represent total oil spilled.
Statistical Issues:	This performance measure is reported by the "spiller," and the probability of a spill being reported is proportional to its size, which implies that the unreported spills are generally the smaller ones. According to USCG, the error rate for volume spilled is less than 5% because larger spills receive the highest levels of case review, and they account for most of the volume spilled.
Verification & Validation:	Verification and validation occurs at several levels. Edit checks within MSIS can detect some incorrect or missing data and force review and correction before data entry is completed. Selection lists for certain data fields also reduce the opportunity for data entry error. All investigations go through one level of review at the field unit for accuracy. Investigations of spills are also usually reviewed at district and headquarters offices. The headquarters Data Administration staff conducts periodic quality control checks to identify entry errors such as missing data or miscoding, and corrects any errors identified. Each spill involving 1000 gallons or more is reviewed before it is included in the measure. Errors identified are referred to either the Data Administration staff or the Investigations and Analysis staff for correction.

Fisheries protection - (FY 2000 – FY 2001)

Page 115

Measure:	Percent change in number of species that are designated as overfished (Includes all areas where Coast Guard has enforcement responsibility in fisheries management plans).
Scope:	This measure includes species covered under formal fisheries management plans that contain Coast Guard enforcement responsibilities, and that are formally assessed by the National Marine Fisheries Service and designated over-fished, approaching over-fished, or not over-fished.
Source:	National Marine Fisheries Service. Data is provided through the annual NMFS report to Congress "Status of Fisheries of the United States." This report is mandated by the Sustainable Fisheries Act of 1996.
Limitations:	Historical data are limited – 1997 - 1999 only. Not all species required to be assessed have been formally assessed as over-fished or not over-fished. Hence, the number of reported over-fished species will likely rise over the next 2 years as reporting becomes more complete. Assessments of over-fishing are based on biological sampling methods and estimations that are subject to error.
Statistical Issues:	As noted in the Limitations section, this measure is likely to rise as NMFS continues its search for currently unknown fish stocks.

Verification & Validation:	Data are provided by NMFS. DOT does not independently verify or validate the data.
Comment:	<p>This measurement is aligned with the Sustainable Fisheries Act and the National Marine Fisheries Service (NMFS) related goal.</p> <p>The Coast Guard also measures the rate of compliance with federal regulations as a critical component of enforcing fisheries management plans designed to improve species health, and prevent over-fishing.</p>

Fisheries protection – (FY 1999)

Page 115

Measure:	Compliance rate with Federal fisheries regulations.
Scope:	This measure includes the observed rate of compliance. It does not include violations that occur, but are not noted by enforcement personnel.
Source:	Internal – Violations and boardings are reported by field units via Coast Guard mission assessment reports. Data is aggregated at the headquarters level.
Limitations:	The observed compliance rate is probably higher than the actual compliance rate because violators actively try to hide non-complying activity. The exact magnitude of this error is unknown. This measure is an intermediate outcome. However, it is a critical measure for gauging the effectiveness of Coast Guard enforcement responsibilities that are necessary to make fisheries management plans viable. It is used in conjunction with the percentage of species designated as over-fished.
Statistical Issues:	As noted in the Limitations section, this measure is an observed compliance, which highlights a potential positive bias in the performance measure.
Verification & Validation:	Primary verification and validation is conducted at the field unit level where the data is entered into mission assessment reports. Validation is also conducted at the district and program manager levels.
Comment:	This is an output measure closely tied to the outcome measure of reducing over-fished stocks. The Coast Guard tracks compliance as one measure of program management.

Hazardous Materials Spills - (FY 2000 – FY 2001)

Page 117

Measure:	Tons of hazardous liquid materials spilled per million ton-miles shipped by pipeline.
Scope:	Hazardous liquid pipeline incidents are those that result in a fatality or injury resulting in hospital treatment or hospitalization property damage equal to or greater than \$50,000, or over 50 barrels spilled. This measure tracks only releases from hazardous liquid pipelines to the environment. Natural gas pipeline releases vaporize into the atmosphere and do not have long-term significant impact on the environment, and thus are not included in this measure.
Source:	Pipeline operators report to RSPA on form 7000-1, Hazardous Liquid Accident Report. RSPA records the data in RSPA's Hazardous Materials Information System.
Limitations:	Because of the magnitude and frequency of fluctuations in the historical data for this measure, a short-term goal will be of limited use in tracking program performance. RSPA does not collect volume shipped data but uses the Association of Oil Pipelines annual Fact Sheet as source for this part of the measure.
Statistical Issues:	These spill incidents are rare and probably not independent events. The performance measure is a ratio, so uncertainty in the denominator can have a large effect on the overall uncertainty.
Verification & Validation:	RSPA reviews the data for accuracy. Supplemental reports are requested where obvious reporting shortcomings are indicated. Additionally, the ASME B31.4 liquid pipeline data review subcommittee performs an annual examination of the hazardous liquid incident reports. Known problems with under-reporting property damages and spill quantities are being addressed by a new industry data improvement effort being piloted in 1999 that will provide better precursor data and more extensive data about impacts to the environment of hazardous liquid pipeline spills.

Comment:	The data for this measure fluctuate year to year. RSPA is studying the spill data to determine the nature of this fluctuation and improve this measure.
----------	---

Hazardous Materials Spills – (FY 1999)

Page 117

Measure:	Gallons of hazardous materials spilled (non-pipeline) per serious transportation incident.
Scope:	Serious reported hazardous materials incidents are defined by RSPA to be those that result in a fatality or major injury (for most purposes, an injury resulting in hospitalization) due to a hazardous material, closure of a major transportation artery or facility or evacuation of six or more persons due to the presence of a hazardous material, or a vehicle accident or derailment resulting in the release of a hazardous material. This measure tracks only transportation related releases of hazardous materials that are in commerce. Volume of spills is not tracked, as this does not necessarily indicate risk.
Source:	Carriers of hazardous materials report to RSPA on DOT Form F 5800.1, Hazardous Materials Incident Report. RSPA records the data in RSPA's Hazardous Materials Information System.
Limitations:	Because of the magnitude and frequency of fluctuations in the historical data for this measure, a short-term goal will be of limited use in tracking program performance.
Statistical Issues:	Self-report data are problematic---assessment of coverage gaps is difficult.
Verification & Validation:	RSPA reviews the data for accuracy. All fatality and injury reports are verified with the reporter as being caused by the hazardous material. Supplemental data are requested where obvious reporting shortcomings are indicated.
Comment:	The data for this measure fluctuates year to year. RSPA is reviewing its definition of serious incidents to better reflect program performance.

DOT facility cleanup

Page 119

Measure:	Percentage of DOT facilities categorized as No Further Remedial Action Planned (NFRAP) under the Superfund Amendments and Reauthorization Act (SARA).
Scope:	EPA maintains a Federal Facility Hazardous Waste docket (docket), which contains information regarding Federal facilities that manage hazardous wastes or from which hazardous substances have been or may be released. DOT facilities listed on the docket are discussed in the Annual SARA report sent to Congress each year. EPA regional offices make the determination to change facility status to NFRAPs on the docket.
Source:	Annual SARA Report to Congress.
Limitations:	The number of DOT facilities listed on the docket can and has fluctuated over the years. Several of the DOT facilities listed have more than one site requiring cleanup and a facility is not removed from the list until all of the sites have no further remedial action planned. Some facilities are listed erroneously and it may take several years to remove them from the docket. NFRAP decisions may be reversed by EPA if future information reveals that additional remedial actions are warranted.
Statistical Issues:	There is no significant error present in the subject data.
Verification & Validation:	The data used in measuring this performance is based on restoration activities at field locations for USCG, FAA, FHWA, and FRA. These field sites report their activities to their respective headquarters management who verifies the data by periodic follow-up reviews. The data is then reported yearly to the Office of the Secretary, who cross-checks it against data received from EPA and the states.
Comment:	The primary criterion for NFRAP is a determination that the facility does not pose a significant threat to the public health or environment. NFRAP decisions may be reversed if future information reveals that additional remedial actions are warranted. The Operating Administrations' activities are controlled, to a degree, by interaction and decisions made by EPA Regional personnel.

Environmental Justice**Page 121**

Measure:	Number of environmental justice complaint cases that remain unresolved after one year.
Scope:	Data will cover complaints filed with DOT under Title VI of the Civil Rights Act of 1964 and which have had environment justice elements, such as allegations of substantially adverse environmental or health impact on a minority or low income community by a transportation project. Case resolutions are actions that end or administratively close out complaints. These include such actions as withdrawals by complainants, resolutions achieved through alternative dispute resolution, findings of no violation, and negotiated settlements after discrimination findings under Title VI.
Source:	DOT will collect this data through the External Complaint Tracking System (XTRAK).
Limitations:	This measure is an initial indicator of how well DOT processes EJ complaints. We anticipate that the number of EJ complaints will be relatively low, compared to other civil rights external complaints. Variables which will not necessarily be assessed include such factors as magnitude of injury, number of beneficiaries adversely affected, pervasiveness, and time constraints before irreparable damage occurs. Other statutory requirements exist for NEPA concerns.
Statistical Issues:	There is no significant error present in the subject data.
Verification & Validation:	Data will cover the entire universe of external complaints received by DOT, and will be entered into the system by operating administrations and DOT Office of Civil Rights staff.
Comment:	This indicator does not measure the impact of DOT's efforts to prevent the conditions that give rise to complaints. It does provide an initial measure of response timeliness, which is important to the public. All environmental justice cases by definition relate to the concerns of a community of low income and/or minority people. In addition, the number of cases indicates the pervasiveness of community perception of significantly adverse environmental and health concerns.

Details on DOT Measures of National Security**Aviation security****Page 127**

Measure:	Detection rate for explosives and weapons that may be brought aboard aircraft.
Scope:	Automated threat-image projection (TIP) and FAA field agent testing of aviation security screener proficiency to detect and resolve images of FAA test objects that simulate deadly or dangerous weapons or explosive devices in checked or carry-on baggage or carried on the person through an airport security checkpoint.
Source:	FAA Office of Aviation Security Airport and Air Carriers Information Reporting System (AAIRS)
Limitations:	No comment.
Statistical Issues:	There is no significant error present in the subject data.
Verification & Validation:	Special "red team" testing led by agents based at FAA headquarters are used to validate automated and field agents' test results.
Comment:	The White House Commission recommended more aggressive, realistic testing which was later required by Public Law 104-264. Funding that began in 1997 enabled an increase in testing as more field agents were hired and trained. Prior to 1998, data from realistic testing was too sparse to be considered conclusive.

Critical infrastructure protection**Page 130**

Measure:	Of those who need to act, percent that receive threat information within 24 hours.
-----------------	---

Scope:	Threat information, in this context, is defined as credible information (both time-sensitive/action-oriented and informational) received by the Intelligence Community, analyzed by the Office of Intelligence and Security (OIS) and distributed in the form of a Transportation Security Information Report, generated by OIS for distribution by the Operating Administrations (OAs). Figure is derived from the percentage of transportation security officials and industry representatives that receive threat information from OIS through the OAs within the 24-hour period. Security representatives and officials will be randomly sampled by OIS within 48 hours of information dissemination and asked if and how soon they received the subject material.
Source:	Internally prepared. Survey conducted by OIS of both DOT personnel and industry security contacts.
Limitations:	Data: Relies on the reporting of the customers and consumers of this information. Reporting could be skewed to reflect positively on the dissemination process within the Operating Administrations. Indicator: This measure only identifies whether there are possible breakdowns and bottlenecks in the dissemination process. It does not identify where those breakdowns specifically may be in the dissemination chain.
Statistical Issues:	Since these data are collected through a sample survey, they are subject to sampling and nonsampling errors.
Verification & Validation:	Customers will be randomly surveyed at all levels within the dissemination process, not solely the end users. Consequently, the reporting of dissemination times and officials who are in receipt of the information can be cross-checked for verification and validity of data.

Sealift capacity

Page 132

Measure:	<ol style="list-style-type: none"> 1. Ship capacity (in twenty-foot container equivalent units, or TEUs) available to meet DOD's requirements for intermodal sealift capacity. 2. Ship capacity (in million square feet) available to meet DOD's requirements for intermodal sealift capacity. (This measure will be discontinued after 2000)
Scope:	Includes the aggregate TEUs (or estimated square footage) of cargo capacity for ships enrolled in the Maritime Security Program (MSP) and Voluntary Sealift Agreement (VISA).
Source:	MARAD/USTRANSCOM database of militarily useful sealift capacity enrolled in the MSP and VISA programs based on vessel capacity data obtained from the vessel operators.
Limitations:	MARAD, DOD and operator data on vessel characteristics (e.g., deck strength in pounds per square feet, deck height, container stowage factors), which are used to determine the portion of a vessel suitable for carrying military cargo, are not always consistent. For example, the majority of ships in MSP/VISA are containerships, which normally are measured in TEUs; however, DOD generally measures surge sealift ships, most of which are Roll-on/Roll-off vessels, in square feet. MARAD is working with DOD and the maritime industry to develop the most accurate data on the portion of each vessel's capacity that is militarily useful. The secondary indicator of square feet will be dropped beginning with FY 2001 until a more accurate indicator is identified by MARAD and DOD. Historical data prior to FY 1997 are unavailable since the MSP and VISA programs were not enacted until that year.
Statistical Issues:	None.
Verification & Validation:	MARAD works with DOD and the maritime industry to use the most accurate information. MARAD validates vessel capacity data, which are obtained from the vessel operators, through comparisons with internationally recognized databases of vessel characteristics (such as Lloyd's Register data), vessel trim and stability information, stowage plans and other cargo loading documents.

Mariner availability**Page 134**

Measure:	Of the mariners needed to crew combined sealift and commercial fleets during national emergencies, the percent of the total that are available.
Scope:	The availability of licensed and unlicensed mariners is determined based on the pool of mariners qualified to crew U.S.-flag oceangoing merchant vessels over 1,000 gross tons that are actively sailing, those that have sailed within five years, and current-year licensed graduates of federally supported maritime academies. This pool is then compared to the DOD and commercial manpower requirements to determine sufficiency of the labor force. The 1999-2001 targets are based on a sealift operation that extends beyond 6-months, necessitating relief for the mariners who were sailing at the start-up of the operation. The actively sailing workforce exceeds the number of full-time shipboard positions because the nature of shipboard employment requires work schedules of 8-12 hours per day, seven days per week during a voyage. Crewmembers are also rotated during the course of the year, with a portion of the workforce at sea while the remainder of the workforce is on shoreside relief, resulting in an average mariner to job ratio of 2.3 for each position.
Source:	Internally prepared: MARAD (crew requirements based on actual crewing levels submitted by U.S-flag operators for the commercial fleet, and by MARAD and DOD for Government-owned organic sealift); Coast Guard Merchant Mariner Licensing and Documentation System (MMLD) merchant mariner sea service data (mariners who have sailed).
Limitations:	<p>The surge in demand for mariners to support an extended mobilization could result in shortages of specific skills required to crew both commercial and DOD organic vessels, even if the aggregate labor pool is sufficient to meet requirements. In an actual mobilization mismatches of the skills of available individuals compared to the skills required for unfilled positions (e.g., a third mate is available to work, but a second mate's position is open) could lead to sporadic crewing shortfalls.</p> <p>Mariner availability is derived by MARAD primarily using data from the MMLD system. The MMLD system provides information on the sea service of licensed and unlicensed mariners. In addition, it records transactions (original issue, renewals, and upgrades) affecting licensed mariners and it maintains grades and ratings of licensed and unlicensed mariners whether or not the seafarers are actively sailing. The Coast Guard recognizes that the MMLD has data integrity problems, including incomplete and incorrect data entries, which preclude direct and comprehensive time-series analysis of licensed mariners. Full-year service data for all operators may not be reported or may lag for several months. Based on data for prior years, approximately 85 percent of FY 1999 MMLD sea service and unlicensed rating records were received by the Coast Guard and provided to MARAD. Full year FY 1999 results were extrapolated based on this proportion. It is uncertain when full year FY 1999 data will become available. The Coast Guard is reengineering the MMLD and additional information is expected to be available for the FY 2000 performance report.</p>
Statistical Issues:	The primary area of uncertainty lies within the MMLD system, which is used to define the population of available mariners. For approximately 20% of the individuals within the historical database, MARAD could not determine whether the original issue transaction was valid. But this problem is less serious for the more recent years.
Verification & Validation:	MARAD intends to verify the number of mariners available by comparing crewing data reported to the Coast Guard with data obtained from the maritime unions. The unions have agreed to provide such data. MARAD will consult with the Coast Guard and the Bureau of Transportation Statistics about the feasibility of using survey techniques to establish the extent of the non-reporting and coverage deficiencies in the MMLD data.
Comment:	The requirement is based on planning factors of 2.3 mariners-per-billet for oceangoing vessels over 1,000 gross tons in the commercial and DOD organic fleets. This reflects average peacetime crew rotation practices and establishes the maximum demand for mariners. In an actual mobilization, however, crew schedules could be extended, which would reduce the number of mariners required.

DOD-designated port facilities**Page 136**

Measure:	Percentage of DOD-designated primary or alternate port facilities that are available when requested by DOD.
-----------------	--

Scope:	The measure consists of the total number of DOD-designated primary ports for military use that are assessed as able to meet DOD-readiness requirements on 48-hour notice, expressed as a percentage of the total number of DOD-designated strategic ports. Port readiness is based on monthly reports submitted by the ports and semi-annual port readiness assessments by MARAD and the Military Traffic Management Command. The MARAD/DOD semi-annual port assessments provide data or other information on a variety of factors, including the following: the capabilities of channels, anchorages, berths, and pilots/tugboats to handle larger ships; rail access, rail restrictions, rail ramp offloading areas, and rail storage capacities; the availability of trained labor gangs and bosses; number and capabilities of available cranes; long-term leases and contracts for the port facility; distances from ports to key military installations; intermodal capabilities for handling containers; highway access; number of port entry gates; available lighting for night operations; and number and capacity of covered storage areas and marshalling areas off the port.
Source:	MARAD data are derived from monthly reports submitted by the strategic ports and from MARAD/DOD semi-annual port assessments.
Limitations:	Port readiness assessments were not made prior to 1995; therefore, data are only available for 1995 and later years. MARAD conducts a monthly survey of all strategic facilities to determine whether they meet the DOD availability requirement. This information is provided to MARAD as a self-assessment by the port agency that owns the facility; there is some degree of subjectivity in determining the availability. As part of the overall planning process, MARAD and DOD conduct semiannual visits to independently verify and reassess port capability and availability. The indicator is by definition a point-in-time judgment. The results of the monthly and semi-annual reports used to measure port readiness can vary in accordance with the intensity of commercial activity at a given port at the time of the assessment. Also, the monthly reports do not include the same level of detail as the semi-annual assessments, although MARAD is in continuous contact with port officials to minimize response error.
Statistical Issues:	The measurement of port readiness is an overall measure derived from MTMC comments, monthly readiness reports and semi-annual assessments. As such, it is a subjective measure.
Verification & Validation:	The MARAD/DOD semi-annual port visits independently verify and reassess not only the DOD-designated facilities, but also the total capability of the strategic port.

Ready Reserve Force (RRF) activation

Page 138

Measure:	<ol style="list-style-type: none"> 1. Percent of RRF no-notice activations that meets assigned readiness timelines. 2. Percent of days that RRF ships are mission-capable while under DOD control.
Scope:	<p>By Memorandum of Agreement with the Department of Defense (DOD), RRF ships are each assigned to 4, 5, 10, 20, or 30 days readiness timeline status, within which stated times they have to be activated and ready to sail. When the Military Sealift Command (MSC) orders the activation of a RRF ship, they precede the hard-copy message with a telephone call in which MARAD and MSC agree on the time that the activation timeline starts to run. MARAD then commences the activation process, which ends when the ship is accepted as Ready for Sea (RFS) by MSC.</p> <p>Once ships are activated and RFS for other than no-notice tests, they are assigned to specific short- or long-term military missions. It is then MARAD's goal to maintain an operational ship in a condition that is capable of fulfilling its mission, including the required sailing speed and the capability for loading and discharging cargo whenever required. Non-performance time is based on the MSC Casualty Reporting (CASREP) system, which identifies casualties that are of a severity to prevent the ship from performing the mission. These messages are passed from the ship's Captain to MSC, the Ship Manager, and MARAD.</p>
Source:	MARAD data on the number of days to activate RRF ships and their operational reliability.
Limitations:	The Army's Warfighting Reserve (AWR-3) program started in FY 1994, and it was at that point that MARAD started to collect operational reliability data. FY 1994-1996 operational reliability data include only the 8 vessels in the Army Prepositioning Stock (APS) program; FY 1997-1999 (as of 6/30/99) data is based upon all activated RRF ships.

Statistical Issues:	Since the population under review consists of 8 vessels from 1994 to 1996 and 15 vessels at preset, large swings in the performance measures are likely due to just one ship being out of service (as was the case in 1997).
Verification & Validation:	<p>The source of the activation data is the actual activation order from DOD to MARAD and the documents produced during a no-notice activation. These fix the actual time of call-up and the time when the vessel is "Ready for Sea"(or tender to MSC). The Ready for Sea time is agreed to by MARAD and the on-board MSC representative and reported to DOD by official message. The time taken to activate each ship is maintained in the ship's logbook and in official DOD, MSC, and MARAD records.</p> <p>The collecting of data regarding mission capability under MSC operational control starts when MSC officially accepts delivery of RRF ships with date and time documentation. All problems that are of a severity to prevent the ship from performing the mission are reported by the Captain of the ship to MSC, the Ship Manager, and MARAD. The ship's Captain also reports when the problem has been corrected. This information is entered by MSC into its CASREP system.</p>
Comment:	The RRF is funded by the Department of the Navy through the National Defense Sealift Fund.

Military Readiness (USCG) - (FY 1999)

Page 140

Measure:	Readiness rating (weighted index) for all high endurance cutters, patrol boats, and port security units
Scope:	Only high and medium endurance cutters, patrol boats, and Port Security Units that are designated as necessary for defense plans are included. The number of specific units required is classified.
Source:	Department of Defense Status of Readiness and Training System (SORTS) – Database used by the Coast Guard in applying DOD standards to its assets to determine a readiness score.
Limitations:	SORTS uses a multi-factor matrix to calculate the readiness status. Ratings are determined through self-assessment by the units. Although specific criteria are outlined for each factor, some judgment is required in applying criteria. Different units and personnel may apply criteria slightly different.
Statistical Issues:	This performance measure is being phased out.
Verification & Validation:	Units self assess and report readiness using objective standards. Unit readiness is periodically validated through inspections, assistance visits, and in some cases training and assessment at Navy facilities. These assessments are conducted by external, field level commands (such as Coast Guard areas, districts, and groups).
Comment:	The readiness rating is determined by a multi-factor matrix that calculates an overall readiness value: C1 is the highest rating, C5 the lowest. These standard, DOD ratings describe the military readiness of a unit "at a point in time." The Coast Guard readiness index is calculated by determining the percentage of units that achieve an average rating of C2 when they are required to be underway or in an on-call status (during planned maintenance periods, units are deliberately placed in a lower readiness status). These percentages are then weighted (0.25 each for high endurance cutters, 0.25 for medium endurance cutters, 0.25 for patrol boats, and 0.25 for PSUs) to arrive at an aggregated index score. The target of 72 accounts for a percentage of new Port Security Units that will not achieve full readiness in 2000.

Military Readiness (Coast Guard) – (FY 2000 – FY 2001)

Page 140

Measure:	Percentage of days that the designated number of critical defense assets (high and medium endurance cutters, patrol boats, and port security units needed to support Defense Department operation plans) maintain a Combat Readiness rating of 2 or better.
Scope:	Only high and medium endurance cutters, patrol boats, and Port Security Units that are designated as necessary for defense plans are included. The specific units required are classified.

Source:	DOD Status of Readiness and Training System (SORTS) – Database used by the Coast Guard in applying DOD standards to its assets determine a readiness score.
Limitations:	SORTS uses a multi-factor matrix to calculate the readiness status. Although specific criteria are outlined for each factor, some judgment is required in applying criteria. Different units and personnel may apply criteria slightly different.
Statistical Issues:	This particular performance measure in FY 1999 is based on two types of boats, patrol boats and endurance cutters, which have extremely different levels of readiness.
Verification & Validation:	Units self assess and report readiness using objective standards. Unit readiness is periodically validated through inspections, assistance visits, and in some cases training and assessment at Navy facilities. These assessments are conducted by external, field level commands (such as Coast Guard areas, districts, and groups).
Comment:	Coast Guard will reassess the overall adequacy this measure.

Drug interdiction

Page 143

Measure:	Seizure rate for cocaine that is shipped through the transit zone.
Scope:	Seizure rate is a measure consisting of the amount of cocaine seized by the Coast Guard divided by the noncommercial maritime cocaine flow, expressed as a percentage. Noncommercial is defined as any vessel or aircraft not engaged in port-to-port transfer of cargo/passengers for the primary purpose of business profit. Examples are pleasure craft, fishing vessels, offshore work-boats, or freighters carrying cargo as a means of disguising illegal drugs.
Source:	The amount of cocaine flow shipped by non-commercial means through the transit zone is estimated in the Interagency Assessment of Cocaine Movement (IACM) published by the Office of National Drug Control Policy (ONDCP). The amount of cocaine seized is measured by Coast Guard crews and reported through the Coast Guard Law Enforcement Information System.
Limitations:	It is probable that non-entry, duplication, and coding errors are present in seizure amount data (numerator), however this error is likely to be small. The cocaine flow amount (denominator) is estimated through a complex process using many different sources of information. Due to the secretive nature of the illegal drug trade, cocaine flow estimates may contain significant errors. The size of this error may vary from year to year; the extent of this is not known. The estimation process changes slightly each year as improvements are made, so year-to-year comparisons of the flow are not completely consistent. The accuracy of the official cocaine flow estimate has been questioned by some individuals and organizations outside of government that have an interest in U.S. drug policy. ONDCP continuously attempts to refine this estimate to improve the measurement of interdiction activities. This measure only addresses cocaine; formal flow assessments do not exist for other major drugs. This measure is not designed to include cocaine shipped by commercial means such as large container vessels; however, it is probable that a small amount of cocaine included in the numerator is actually related to commercial shipping. This distinction between commercial and noncommercial is better for program management; at-sea interdiction of commercially conveyed cocaine, particularly when shipped in containers is extremely difficult, and not the focus of the Coast Guard program.
Statistical Issues:	The primary source of uncertainty in estimating seizure rate for cocaine is the denominator. Although the numerator estimate of cocaine seized is relatively accurate, the estimate of the amount shipped in the denominator is far more variable.
Verification & Validation:	Verification and validation occurs in several places in the data reporting and collection process. Data entry software helps ensure data quality and consistency by employing selection lists and logic checks. Internal analysis and review of published data by external parties help identify errors.
Comment:	This measure is consistent with the goals contained in the President's National Drug Control Strategy.

Migrant interdiction

Page 145

Measure:	Success Rate for undocumented migrants attempting to enter the U.S. over maritime routes.
-----------------	--

Scope:	Measure includes Cuban, Dominican, Haitian, and Chinese migrants, as these are the primary groups using maritime channels and the groups for which formal flow estimates exist. Success rate is the estimated number arriving by maritime channels divided by those that pose a threat of migration (actual interdictions plus estimated intent).
Source:	Data is obtained from Coast Guard and from the Immigration and Naturalization Service (INS). Actual interdiction numbers come from direct counts by Coast Guard, U.S. Border Patrol, and other official sources. Estimates of migrants who successfully arrive and estimates of those with a high potential for undertaking the voyage are derived (with a consistent methodology) from investigations of incidents, interviews of detainees, and intelligence gathering. Sources for this information are the Coast Guard, INS, and other authorities.
Limitations:	The numbers of illegal migrants entering the U.S., and the numbers of potential migrants, are derived numbers subject to estimating error. Because of the speculative nature of the information used, and the secretive nature of illegal migration, particularly where professional smuggling organizations are involved, the estimated potential flow of migrants may contain significant error. The measure only tracks four migrant groups at this time. A small number of migrants (approximately 10%) from various source countries are not included because formal flow estimates of migrants leaving these countries is not available. Using the number of potential migrants in the denominator helps address the deterrence value of Coast Guard operations, but could lead to confusion of this measure with a simple interdiction rate. Trend information for 1995-1998 is available.
Statistical Issues:	The primary source of uncertainty in estimating the success rate for undocumented migrants is the denominator, which is an estimate of the flow of migrants, both documented and undocumented.
Verification & Validation:	The numbers of migrants reaching the U.S. via maritime routes and the number of "potential" migrants" are estimated. Methodologies and data are continuously reviewed. The Coast Guard has developed the estimation techniques that support this indicator over the last five years in order to more consistently use intelligence information. They are seeking independent assessment of the methods, and look to improve the process in the future.
Comment:	Partly because maritime threats of illegal migration have come from a limited number of sources, the Coast Guard and others have developed quantified threat estimates to better manage interdiction. Over the past five years, estimates have been formalized in a process that removes as much subjectivity and inconsistency as possible. It should be noted that past information reflects the success of intentional illegal activity. In contrast to some DOT measures, future program outcomes can not be confidently projected from measures of the immediate past.